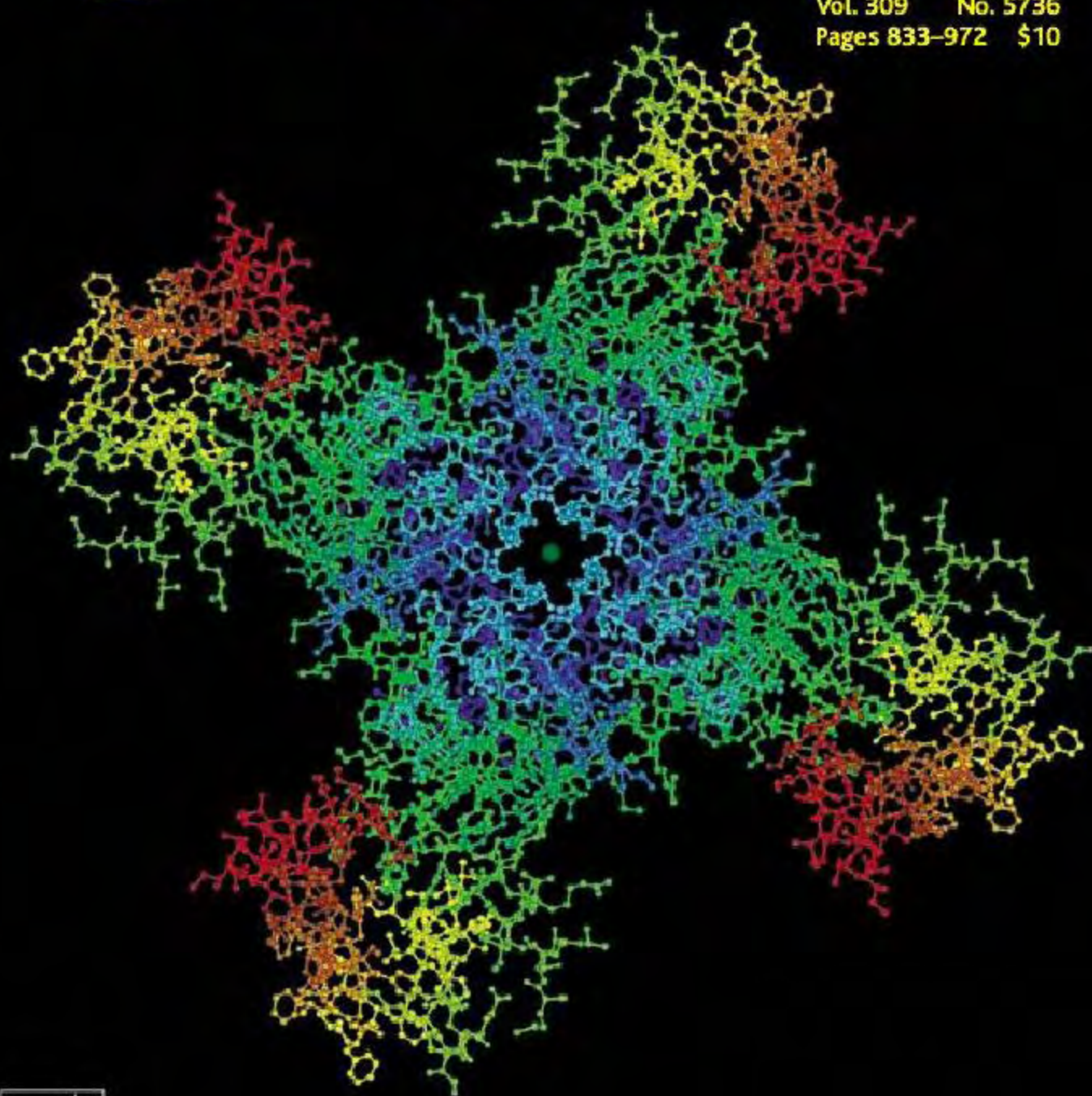


Science

5 August 2005

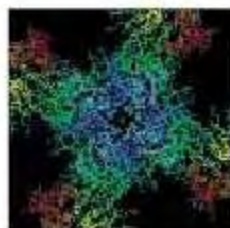
Vol. 309 No. 5736

Pages 833-972 \$10



125
YEARS OF GLOBAL
Science

AAAS



COVER The integral membrane portion of the mammalian Shaker family potassium channel, Kv1.2, viewed from the extracellular solution. The channel is gradually colored from its amino terminus (red) to its carboxyl terminus (blue) and is illustrated with sticks. A green sphere represents a potassium ion in the central ion conduction pathway. See page 897. [Image: S. B. Long et al.]

DEPARTMENTS

- 845 **SCIENCE ONLINE**
- 847 **THIS WEEK IN SCIENCE**
- 851 **EDITORIAL** by Jerome F. Strauss III
NIH Funding Reform
- 852 **EDITORS' CHOICE**
- 856 **CONTACT SCIENCE**
- 857 **NETWATCH**
- 955 **NEW PRODUCTS**
- 956 **SCIENCE CAREERS**

NEWS OF THE WEEK

- 858 **STEM CELL POLITICS**
Frist's Support Raises Odds for Passage of Stem Cell Bill
- 859 **PLANETARY SCIENCE**
Cassini Catches Mysterious Hot Spot on Icy-Cold Enceladus
- 859 **PLANETARY SCIENCE**
Newfound 'Tenth Planet' Puts Pluto Behind the Eight Ball
- 860 **AIDS RESEARCH**
Male Circumcision Thwarts HIV Infection
- 861 **PALAEANTHROPOLOGY**
U.S. Government Shifts Stance on Claims to Ancient Remains
- 861 **SCIENCE SCOPE**
- 862 **BIOCHEMISTRY**
Cut-Rate Genomes on the Horizon?
related Science Express Report by J. Shendure et al.
- 862 **CLONING**
The Perfect Pedigree
- 863 **ENERGY POLICY**
U.S. Energy Bill Promises Some Boosts for Research

NEWS FOCUS

- 864 **NEUROSCIENCE**
Preventing Alzheimer's:
A Lifelong Commitment?
- 867 **NEUROSCIENCE**
A New Portrait Puts Potassium Pore in a Fresh Light
related Research Articles pages 897 and 903
- 868 **MEETING**
Rencontre Assyriologique Internationale
Alas, Babylon: Tracing the Last King's Desert Exile
Ur's Xena: A Warrior Princess for Sumeria?
Looted Tablets Pose Scholar's Dilemma
- 870 **EPIDEMIOLOGY**
Drugs, Quarantine Might Stop a Pandemic Before It Starts
related Science Express Report by L. M. Longini Jr. et al.
A Drug Makes It Big—But Can It Deliver?



864



880



890

- 872 **RANDOM SAMPLES**

LETTERS

- 874 **Scrapie in Ancient China?** R. B. Wickner; **Co-Funding in Canada: Another View** A. J. Carty; **Update on the Closure of a Chilean Institute** Y. Provoste and E. Gales; **Canadian Database in Singapore** C. Hogue; **Random Copying and Cultural Evolution** R. A. Bentley and S. J. Shennan; **An Unexpected Cover Image** L. Felber
- 879 **Corrections and Clarifications**

BOOKS ET AL.

- 880 **EVOLUTION**
Evolution of the Insects
D. Grimaldi and M. S. Engel, reviewed by E. Jarzembowski
- 881 **NATURAL HISTORY**
Robbing the Bees: A Biography of Honey: The Sweet Liquid Gold that Seduced the World: Letters from the Hive: An Intimate History of Bees, Honey, and Humankind: The Hive: The Story of the Honeybee and Us
H. Bishop; S. Buchmann and B. Reppner; B. Wilson; reviewed by J. Hostler
- 882 **Browsings**

ESSAY

- 883 **GLOBAL VOICES OF SCIENCE**
Of Stones and Health: Medical Geology in Sri Lanka
C. Dissanayake



PERSPECTIVES

- 886 **CANCER**
Suppressing Cancer: The Importance of Being Senescent
J. Campisi
- 887 **APPLIED PHYSICS**
Droplet Control for Microfluidics
M. Joanicot and A. Ajdari
- 888 **PLANETARY SCIENCE**
The Enigma of the Martian Soil
A. Barin
- 890 **DEVELOPMENTAL BIOLOGY**
Rac1 Up for Epidermal Stem Cells
G. P. Dotto and G. Cotsarelis
related Report: page 933
- REVIEW**
- 892 **SOCIOLOGY**
The Coming Paradigm Shift in Forensic Identification Science
M. J. Saks and J. J. Koehler

SCIENCE EXPRESS www.sciencexpress.org

EPIDEMIOLOGY: Containing Pandemic Influenza at the Source

I. M. Longini Jr., A. Nizam, S. Xu, K. Ungchusak, W. Hanshaworakul, D. A. T. Cummings, M. E. Halloran
A model of a southeast Asian population predicts that a hypothetical emergent flu strain may be containable with antiviral agents, quarantine, and vaccination. *related News story page 870*

BIOCHEMISTRY: Accurate Multiplex Polony Sequencing of an Evolved Bacterial Genome

J. Shendure, G. J. Porreca, N. B. Reppas, X. Lin, J. P. McCutcheon, A. M. Rosenbaum, M. D. Wang, K. Zhang, R. D. Mitra, G. M. Church
DNA can be sequenced routinely at about one-tenth the cost of conventional sequencing with off-the-shelf instruments and reagents. *related News story page 862*

MOLECULAR BIOLOGY: Inhibition of Translational Initiation by let-7 MicroRNA in Human Cells

R. S. Pillai, S. N. Bhattacharyya, C. G. Artus, T. Zoller, N. Cougot, E. Basyuk, E. Bertrand, W. Filipowicz
A human microRNA regulates gene expression by inhibiting translation initiation, probably by binding to the cap structure at the 5' end of the targeted messenger RNA.

CHEMISTRY: Ultrafast Dynamics of Solute-Solvent Complexation Observed at Thermal Equilibrium in Real Time

J. Zheng, K. Kwak, J. Asbury, X. Chen, I. R. Piletic, M. D. Fayer
Vibrational echo correlation spectroscopy can image the association and dissociation of phenol-benzene complexes over a few picoseconds, a time regime that has been inaccessible to NMR spectroscopy.

BREVIA

896 OCEAN SCIENCE: Extreme Waves Under Hurricane Ivan

D. W. Wang, D. A. Mitchell, W. J. Teague, E. Jarosz, M. S. Hulbert
Tide gauges in the Gulf of Mexico show that in 2004 Hurricane Ivan generated waves over the continental shelf with crest-to-trough heights that may have exceeded 40 meters.

RESEARCH ARTICLES

NEUROSCIENCE

897 Crystal Structure of a Mammalian Voltage-Dependent Shaker Family K⁺ Channel

S. B. Long, E. B. Campbell, R. MacKinnon

903 Voltage Sensor of Kv1.2: Structural Basis of Electromechanical Coupling

S. B. Long, E. B. Campbell, R. MacKinnon

An x-ray crystal structure of a eukaryotic voltage-gated potassium channel, probably in its native conformation, reveals how movement of the voltage sensor triggers opening of the pore. *related News story page 867*

REPORTS

909 MATERIALS SCIENCE: Triangular and Fibonacci Number Patterns Driven by Stress on Core/Shell Microstructures

C. Li, X. Zhang, Z. Cao

When stressed by cooling, a silica coating on silver nanoparticles forms complex, organized patterns similar to those seen in flowers and seeds.

911 MATERIALS SCIENCE: A Crossover in the Mechanical Response of Nanocrystalline Ceramics

I. Szlufarska, A. Nakano, P. Vashita

Simulations show that silicon carbide with nanometer-sized grains deforms first by cooperative slip along soft grain boundaries and then by loss of crystallinity.

914 CHEMISTRY: Characterization of Excess Electrons in Water-Cluster Anions by Quantum Simulations

L. Turi, W.-S. Sheu, P. J. Rossky

Simulations suggest that in anionic water clusters of fewer than 200 molecules, the excess electron resides on the surface of the cluster rather than inside it.

917 CHEMISTRY: Probing the Ultrafast Charge Translocation of Photoexcited Retinal in Bacteriorhodopsin

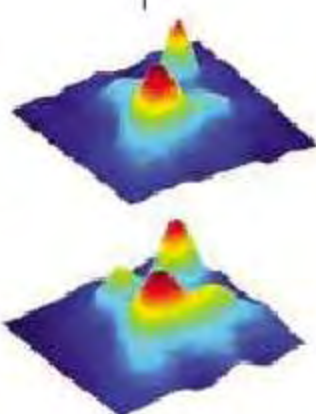
S. Schenkl, F. van Mourik, G. van der Zwan, S. Haacke, M. Chergui

Photoexcitation of the light-sensitive pigment retinal in bacteriorhodopsin increases the dipole moment within 200 femtoseconds, probably driving the subsequent isomerization.

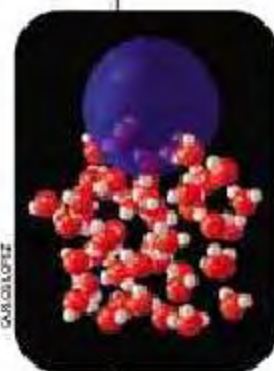
920 FLUID DYNAMICS: An Experimental Approach to the Percolation of Sticky Nanotubes

B. Vigolo, C. Coulon, M. Maugey, C. Zakri, P. Poulin

Nanorods in solution can form a coherent, connected network at much lower concentrations when an added surfactant ensures only weak interactions between them.



896



914

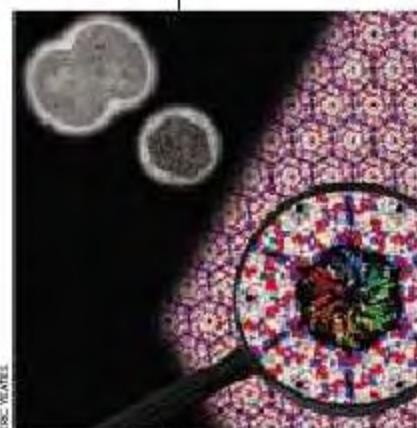
Contents continued ►

REPORTS CONTINUED

- 923 GEOCHEMISTRY:** The Pyrite-Type High-Pressure Form of Silica
Y. Kuwayama, K. Hirose, N. Sata, Y. Ohishi
 Experiments confirm that silica can exist in a dense, high-pressure phase in which each silicon atom is coordinated to six nearby oxygens and two more distant ones.
- 925 CLIMATE CHANGE:** Ice Sheet and Solid Earth Influences on Far-Field Sea-Level Histories
S. E. Baett, G. A. Milne, J. X. Mitrovica, P. U. Clark
 A model with a stiff lower mantle and rapid melting of Antarctic ice sheets matches well the rise in sea level after the last glacial maximum observed at tropical Pacific sites.
- 929 PLANT SCIENCE:** Antagonistic Control of Disease Resistance Protein Stability in the Plant Immune System
B. F. Holt III, Y. Belkhaoui, J. L. Dangl
 Two plant proteins thought to trigger protective pathways upon pathogen attack actually form a regulatory system that keeps defense proteins available for rapid deployment.
- 933 DEVELOPMENTAL BIOLOGY:** Stem Cell Depletion Through Epidermal Deletion of Rac1
S. A. Benitah, M. Frye, M. Glogauer, F. M. Watt
 A small GTP-binding regulatory protein is required for maintaining stem cells in the skin and preventing their differentiation into other epidermal cell types. *related Perspective page 830*
- 936 MICROBIOLOGY:** Protein Structures Forming the Shell of Primitive Bacterial Organelles
C. A. Kerfeld, M. R. Sawaya, S. Tanaka, C. V. Nguyen, M. Phillips, M. Beeby, T. O. Yeates
 The carboxysome, a CO₂-fixing microcompartment in certain bacteria, resembles a viral capsid of hexameric protein building blocks, with pores that may regulate metabolite flow.
- 938 MOLECULAR BIOLOGY:** Rewiring of the Yeast Transcriptional Network Through the Evolution of Motif Usage
J. Ihmels, S. Bergmann, M. Gerami-Nejad, I. Yanai, M. McClellan, J. Berman, N. Barkai
 Yeast species that grow aerobically have a common sequence in the promoters of mitochondrial ribosomal proteins, apparently acquired by a common ancestor.
- 941 MICROBIOLOGY:** Export-Mediated Assembly of Mycobacterial Glycoproteins Parallels Eukaryotic Pathways
B. C. VanderVen, J. D. Harder, D. C. Crick, J. T. Belisle
 In bacteria, sugar residues are added to proteins during export by a mechanism similar to that used by eukaryotes.
- 943 MEDICINE:** Regulation of Blood Glucose by Hypothalamic Pyruvate Metabolism
T. K. T. Lam, R. Gutierrez-Juarez, A. Pocai, L. Rossetti
 A region at the base of the brain functions as the body's glucose monitor, instructing the liver to shut down glucose production when blood glucose levels get too high.
- 948 NEUROSCIENCE:** Hemodynamic Signals Correlate Tightly with Synchronized Gamma Oscillations
J. Niessing, B. Ebisch, K. E. Schmidt, M. Niessing, W. Singer, R. A. W. Galuske
 In cat visual cortex, brain imaging signals correlate more closely with synchronous synaptic activity than with the rate of action potential firing.
- 951 NEUROSCIENCE:** Coupling Between Neuronal Firing, Field Potentials, and fMRI in Human Auditory Cortex
R. Mukamel, H. Gelbard, A. Arieli, U. Hasson, I. Fried, R. Malach
 When a subject is viewing and listening to a movie, brain imaging of the auditory cortex provides a good indication of the underlying neuronal activity.



929



936



ADVANCING SCIENCE. SERVING SOCIETY

Change of address after 4 weeks, if you did not receive address and 5-day notice. Number Postmaster: Send change of address to Science, P.O. Box 1811, Danbury, CT 06810-1811. Single copy sales: \$10.00 per issue. Periodical postage paid at Washington, DC. Publication No. 484460. Periodicals postage paid at Washington, DC, and additional mailing offices. Copyright © 2005 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic institutional subscription (12 issues): \$705 (\$74 allocated to subscription). Domestic institutional subscription (6 issues): \$352.50. Foreign postage extra (Mexico, Caribbean, and other countries): \$55.00. Other countries (air mail delivery): \$85.00. First class, airmail, student, and individual rates on request. Circulation with GST available upon request. GST # R123456789. Publication Mail Agreement Number 130608R. Printed in the USA.

Change of address after 4 weeks, if you did not receive address and 5-day notice. Number Postmaster: Send change of address to Science, P.O. Box 1811, Danbury, CT 06810-1811. Single copy sales: \$10.00 per issue. Periodical postage paid at Washington, DC. Publication No. 484460. Periodicals postage paid at Washington, DC, and additional mailing offices. Copyright © 2005 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic institutional subscription (12 issues): \$705 (\$74 allocated to subscription). Domestic institutional subscription (6 issues): \$352.50. Foreign postage extra (Mexico, Caribbean, and other countries): \$55.00. Other countries (air mail delivery): \$85.00. First class, airmail, student, and individual rates on request. Circulation with GST available upon request. GST # R123456789. Publication Mail Agreement Number 130608R. Printed in the USA.

Contents continued

Humans Drove Giant Sloths to Extinction

New study argues climate had little to do with great North American mammal die-off.

Pollen Packs a Powerful Punch

Newly discovered component of pollen grains is critical for allergic reactions.

An e-Tag for Every Bag

New modifications may someday make electronic chips as ubiquitous as bar codes.



Opportunities for research in fresh water.

science's next wave www.nextwave.org CAREER RESOURCES FOR YOUNG SCIENTISTS

CANADA: Tapping into Freshwater Science A. Fazekas

Next Wave explores the opportunities that exist to work on freshwater issues.

MSciNet: Fulfilling the Expectation of Excellence C. Parks

The Meyerhoff Scholarship Program at the University of Maryland, Baltimore County, has attracted top minority students.

MSciNet: Serving the Native American Community E. Francisco

Alexander Red Eagle hopes to help his community by becoming a biomedical researcher and physician.

POSTDOC NETWORK: Making It Great for Everybody B. Benderly

Efforts to improve conditions for postdocs at the departmental level can make a big difference.

GRANTSNET: August 2005 Funding News Edited by S. Martin

Get the latest index of research funding, scholarships, fellowships, and internships.

science's sage ke www.sageke.org SCIENCE OF AGING KNOWLEDGE ENVIRONMENT

PERSPECTIVE: From Bedside to Bench—Research Agenda for Frailty L. P. Fried, E. C. Hadley,

J. D. Walston, A. Newman, J. M. Guralnik, S. Studenski, T. B. Harris, W. B. Ershler, L. Ferrucci

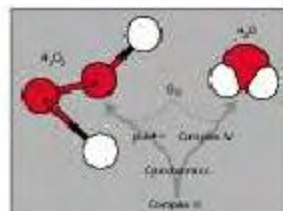
Conference aims for a better understanding of the physiology and etiology of this condition.

News Focus: Not a Chip Off the Old Block M. Leslie

Study identifies unexpected function for mammalian version of yeast longevity protein.

News Focus: Detour to Death R. J. Davenport

Protein kills cells by diverting electrons and crafting free radicals.



Making poisons instead of power.



Structural insight into 14-3-3.

science's stke www.stke.org SIGNAL TRANSDUCTION KNOWLEDGE ENVIRONMENT

REVIEW: 14-3-3 Proteins—A Number of Functions for a Numbered Protein D. Bridges and

G. B. G. Moorhead

14-3-3 proteins can initiate conformational changes in their targets, occlude structural features, or serve as scaffolding.

REVIEW: Histone Deacetylases as Transcriptional Activators? Role Reversal in Inducible Gene Regulation I. Nusinzon and C. M. Horvath

Histone deacetylases modify chromatin and transcription factors to activate as well as repress gene transcription.

Separate individual or institutional subscriptions to these products may be required for full-text access.

GrantsNet
www.grantsnet.org
RESEARCH FUNDING DATABASE

AIDScience
www.aidscience.com
HIV PREVENTION & VACCINE RESEARCH

Members Only!
www.AAASMember.org
AAAS ONLINE COMMUNITY

Functional Genomics
www.sciencegenomics.org
NEWS, RESEARCH, RESOURCES

A Forensic Analysis

The legal system frequently faces situations in which scientifically valid data would help determine the outcome of the case. **Saks and Koehler** (p. 892) review the state of forensic science and find it to be in transition. Some areas, such as DNA fingerprinting, are increasingly well grounded in scientific principles, whereas other areas are more subjective. The authors discuss the various sources of error and offer some proposals for improving the rigor of forensic science.

Flex and Rise

Earth models that have attempted to simulate the sea-level rise from ice sheet melting after the Last Glacial Maximum have failed to reproduce the changes recorded at the so-called "far-field" sites, such as Tahiti and the Sunda Shelf. **Bassett et al.** (p. 925, published online 23 June 2005) have used a model that combines a high-viscosity lower mantle and a significant contribution from the Antarctic ice sheet to meltwater formation. The reconstructed record and the data agree well, and these results also provide another line of evidence that Antarctic ice was responsible for more of the deglacial sea-level rise than was thought until recently.

Softer at the Edges

Metals become harder as grain sizes decrease, but at some point the grains become so small that the deformation mechanisms change. Nanostructured ceramics also show enhanced properties relative to their coarser-grained counterparts, but do similar changes in deformation mechanisms occur in these more brittle materials? **Szulafarska et al.** (p. 911) show that these ceramics can be thought of as composites of hard nanoscale grains bounded by softer, amorphous-like grain boundaries. A massive molecular dynamics simulation shows that nanoindentation of a nanostructured silicon carbide goes through four deformation regimes. The deformation changes from cooperative grain sliding to a process dominated by amorphization of the crystalline grains.

Patterns of Stress

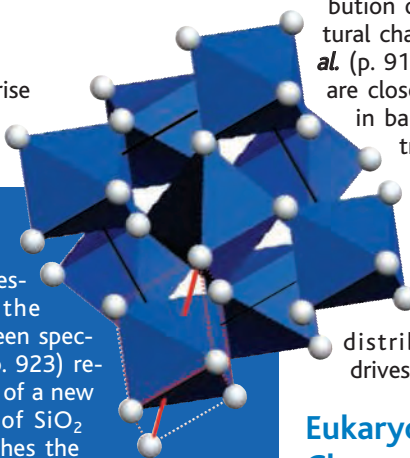
During the fabrication of nanoparticles consisting of a silver core surrounded by a silica shell, **Li et al.** (p. 909) found that controlling the cooling rate could induce stresses in the silica that cause it to form a dimpled pattern on the core sphere. The silica bumps take on either a tri-



angular or Fibonacci sequence pattern that minimizes the total strain energy. These patterns are highly reminiscent of those seen in the development of flowers and plants.

A Little Light Work

Light-driven structural changes in proteins that are required for function are likely the result of photoexcitation processes redistributing charges. However, measuring changes in charge distribution on the time scale of the structural changes is challenging. **Schenkl et al.** (p. 917) have used Trp residues that are close to the retinal-binding pocket in bacteriorhodopsin to probe electric field changes. From the observed changes in Trp absorbance, they calculate that the retinal dipole moment increases during the first 200 femtoseconds after excitation. This change in charge distribution precedes, and likely drives, isomerization.



Squeezed Silica

The existence of a high-pressure form of silica with the pyrite structure has long been speculated. **Kuwayama et al.** (p. 923) report experimental evidence of a new high-pressure polymorph of SiO_2 with a structure that matches the theoretical predictions. Although it is unlikely that this polymorph plays a role in the core of the Earth, this structure has implications for the existence of SiO_2 in the deep planetary interiors of gas giants such as Uranus and Neptune.

Eukaryotic Potassium Channel Structure

Voltage-gated K^+ channels open in response to cell depolarization, reacting to the change in potential by movement of four charged Arg residues, which opens the pore and allows only K^+ ions to exit the cell. X-ray crystallographic structures of bacterial channels have revealed the basis of the K^+ selectivity. Forming crystals of the larger, multisubunit eukaryotic K^+ channels has been more challenging, but **Long et al.** (pp. 897 and 903, published online 7 July 2005; see the cover and the news story by **Service**) now present in two papers a 2.9-angstrom-resolution crystal structure and a mechanistic analysis for eukaryotic Kv1.2 channels from the Shaker family. The crystals, which were formed by adding lipids during crystallization, include the oxido-reductase β subunit and are probably in a native, open state. The β subunits are positioned directly below the intracellular opening to the pore but far enough away to allow the K^+ ions access to the pore through four large side portals. The voltage-sensor domains act as almost independent regions positioned within the membrane beside the cylindrical pore, with at least one of the charge-sensing arginines in direct contact with lipid. Movement of the voltage sensor causes pore opening through the S4-S5 linker helix, which constricts and dilates the S6 "inner" helices around the pore. This structure explains many apparently contradictory results reported to date on K^+ channel structure and function.

Avoiding Too Much of a Good Thing

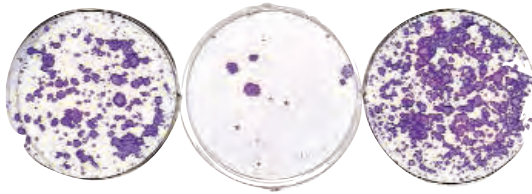
Certain plants carry resistance (R) genes variants that match a particular pathogen's virulence factor. However, too much or too little of the R protein can send the plant's immune response hay-

CONTINUED ON PAGE 849

wire. **Holt et al.** (p. 929, published online 23 June 2005) now provide a genetic analysis of some of the factors that keep the immune response in *Arabidopsis* primed for a rapid deployment but not running rampant. One component, RAR1, somehow promotes the accumulation of the R proteins, and another, SGT1, interacts with RAR1, antagonizing its activity. SGT1 does double duty in infected plants by regulating the cell death response that limits the damage done by some pathogens.

Differentiation on the Rac

Rac1, a member of the Rho family of guanosine triphosphatases (GTPases), is a pleiotropic regulator of many cellular processes, including the cell cycle, cell-cell adhesion, and motility, as well as a key regulator of epithelial differentiation. **Aznar Benitah et al.** (p. 933; see the Perspective by **Dotto and Cotsarelis**) show that Rac1 is expressed



in the pro-liferative compartment of mammalian epidermis. In mice, conditional deletion of Rac1 produces a rapid transient proliferation of cells, followed by the depletion of epidermal stem cells and by a corresponding increase in cell differentiation. For its effect on the stem cell

compartment, Rac1 acts through negative regulation of c-Myc. Thus, as Rac1 is down-regulated, cells can no longer adhere tightly to the substratum, which leads to an inefficient relay of signals from the stem cell niche and subsequent cell differentiation.

Sweet Relations

Bacteria can glycosylate proteins, but the mechanisms and spatial localization of glycosylation are poorly understood relative to those of eukaryotes. **VanderVen et al.** (p. 941) now describe a direct link between prokaryote protein glycosylation and the Sec translocation system, the primary protein export mechanism in bacteria. The association of protein O-mannosylation and Sec-translocation along with other known aspects of protein glycosylation in *Mycobacterium tuberculosis* present parallels with the O-mannosylation system of eukaryotes, in particular the well-studied protein mannosyltransferase system found in budding yeast. Thus, primitive prokaryotes have systems for O-protein glycosylation that are analogous to those present in eukaryotes.

The Ultimate Glucose Monitor

The brain, and in particular, the hypothalamus, controls liver glucose production, but the cellular and molecular mechanisms by which the brain senses glucose levels have been unclear. **Lam et al.** (p. 943) now show that, in rats, this process requires the conversion of glucose in the hypothalamus to lactate, which in turn stimulates pyruvate metabolism and adenosine triphosphate (ATP) production. Alterations in ATP levels control neuronal excitability through effects on ATP-sensitive potassium channels, which have been implicated in glucose output by the liver.

Neuronal Oscillations and Brain Imaging

Brain-imaging methods detect neuronal activity indirectly by measuring blood oxygenation level-dependent (BOLD) signals. **Niessing et al.** (p. 948) investigated the hemodynamic responses recorded by optical imaging and compared them to neuronal activity recorded with microelectrodes in anesthetized cats. The BOLD response was well correlated with the gamma-frequency components of the local field potential, but only weakly correlated with firing rate. **Mukamel et al.** (p. 951) compared electrophysiological measurements obtained from the auditory cortex of neurosurgical patients with functional magnetic resonance imaging (fMRI) signals obtained from the conscious human brains under identical sensory stimulation. A long-lasting coupling was observed between fMRI measurements and single unit activity. Thus, the fMRI signal reflects the firing rate of human cortical neurons during complex natural stimulation.

NIH Funding Reform

Program officers at the U.S. National Institutes of Health (NIH) have been informing applicants either that their grants will not be funded or that their budgets will be slashed to keep paylines from sinking further. Everyone hoped for a “soft landing” after the NIH budget doubled (from 1999 to 2003), but it is clear that the landing more closely resembles a controlled crash. NIH was responsive to Congress in creating new initiatives during that time, but those initiatives were rarely supported by additional appropriations. Ongoing commitments to those programs leave most NIH institutes with little room to refresh research portfolios, shaking the confidence of the extramural research community. Add to this the NIH reauthorization legislation now under discussion, which contains provisions for sweeping change that could cause further chaos, depending on its final configuration. There is already concern that the proposed restructuring of the agency may lead to reallocation of funds that could threaten programs that target specific diseases as well as basic research.

Change is necessary and can be good, but there must be consideration of the collateral damage it can cause. A case in point being the many young physicians emerging from training to do translational research, part of the NIH “roadmap” for biomedical research in the 21st century. They are meant to rebuild the human capital that was decimated in the late 1980s and early 1990s when funding difficulties drove many out of research. The current cadre of physician scholars will emerge, after up to 5 years of training, to find research grants scarce. If they abandon research, an investment of half a million dollars per career development award, in addition to money allocated for a federal loan repayment program, will have been wasted. Universities and medical centers that have leveraged their finances to build infrastructure will also suffer. And NIH itself will be another victim, because extramural investigators must deal with drastic budget cuts that will impair productivity. The ultimate victim will be the U.S. public, who will not realize the full potential of their tax dollar investment.

The hard landing was inevitable and should not have surprised anyone. The roller-coaster nature of appropriations to NIH and the need to expend allocations fully in a given fiscal year were bound to expose flaws in the way that the government sponsors research. The planning process used by NIH institutes in awarding grants, including the way it handles the bolus of amended proposals accumulating in the pipeline, contributes to the seriousness of the problem. The concomitant increase in the number of grant applications submitted and the reorganization of the review groups that assess them, have compounded the problem.

The key issues facing the biomedical research community, NIH, Congress, and the public are how to repair the immediate damage and prevent poorly conceived reactionary change before it gets out of hand. Congress, in considering NIH reauthorization, should recognize that long-term appropriations for NIH, or flexibility in carrying forward uncommitted funds into a national research trust, would provide much-needed stability. Fortunately, NIH is contemplating funding plans that are based on the appropriation horizon and maintain leeway for opportunistic investment. Critical periodic review of long-standing programs and strategies to synchronize training programs with the future availability of research funds are essential. Requests for grant applications should be limited. And temporarily suspending annual cost-of-living increases for funded grants could instantaneously free up dollars without a great impact on the progress of the science.

Congress should avoid burdening NIH with unfunded mandates. Although several are worthy initiatives, they are potential sinkholes that draw funds away from mission-oriented programs. NIH should push back if special-interest proponents are not willing to put up money, and scientists need to inform advocacy organizations that the system may be damaged as a whole if appropriations do not accompany authorization. The positive message is that those who raise money outside the federal system now have a unique opportunity to influence the national research agenda.

As for reauthorizing NIH, with a proposed \$29.4 billion budget in 2006, Congress would be wise to carefully draft legislation that coordinates research efforts in a manner that maintains a steady infusion of grants to medical schools and research institutes who have already adjusted their programs in response to the NIH research portfolio. A critical opportunity is at hand to create conditions that allow the full potential of NIH research to be realized.

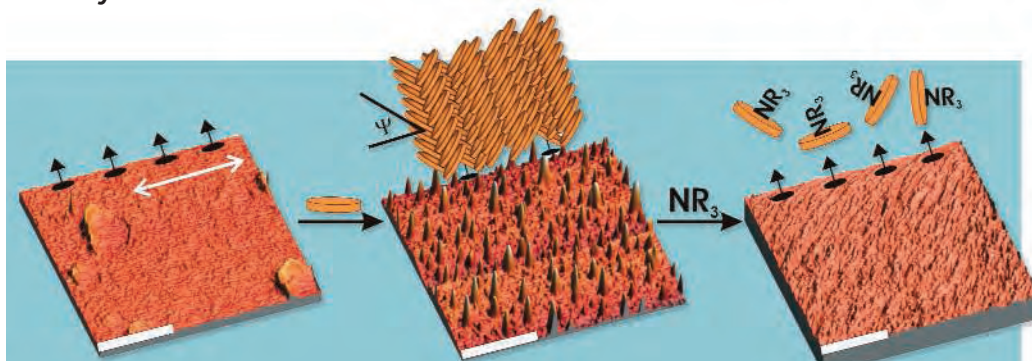
Jerome F. Strauss III

Jerome F. Strauss III is professor and associate chair of the Department of Obstetrics and Gynecology and director of the Center for Research on Reproduction and Women's Health at the University of Pennsylvania Medical Center, Philadelphia, PA.

10.1126/science.1111874



edited by Gilbert Chin



CHEMISTRY

Turn On, Tune In, Drop Out

In a liquid crystal display, the brightness or color of a pixel is controlled by the orientation of the molecules and can be changed by turning on an electric field. In order to ensure uniform alignment within a domain, the glass surface is treated to make it grooved or otherwise anisotropic, which lowers the overall surface energy between the liquid crystal molecules and the glass.

Having shown previously that an oligosiloxane compound deposited onto glass covered with indium tin oxide (ITO) spontaneously forms rigid oligomers that amplify the grooves in the ITO coating, Hoogboom *et al.* use a pyridine-functionalized siloxane that forms an alignment layer capable of binding to the dye zinc phthalocyanine (ZnPc). The ZnPc molecules form epitaxial stacks whose height can be controlled by varying the immersion time, thus providing an opportunity to tune the sizes of domains. After liquid crystal molecules are deposited, it is generally difficult to alter the strength of the surface interactions or to correct defects. However, in this system, adding nitrogen-containing compounds partially dissolves the ZnPc stacks, which drop out and hence provide a second chance to tune the device. — MSL

J. Am. Chem. Soc. 10.1021/ja051865l (2005).

of the effect of black carbon aerosols cannot be evaluated precisely, and black carbon can influence the radiative properties of Earth in other ways, such as by decreasing the albedo of snow. — HJS

Geophys. Res. Lett. 32, 10.1029/2005GL023370 (2005).

BIOCHEMISTRY

Who's the Most Proficient of All?

It might seem a straightforward matter to determine how much the enzyme urease accelerates the rate of decomposition of urea, which, after all, contains only four nonhydrogen atoms. The problem, however, is that in the nonenzymatic pathway, the elimination of ammonia precedes the addition of water, whereas the enzyme promotes the nucleophilic attack of water, yielding a tetrahedral intermediate and a different reaction pathway to the same products.

Previously, Estiu and Merz carried out a computational analysis of the uncatalyzed reaction, relying both on earlier studies with small dinickel molecules and on structural analysis of the dinickel cluster at the urease active site. They found that the catalytic

GENETICS

Luck of the Draw

Genetically identical organisms that have been raised in identical environments age at different rates, suggesting that in addition to genes and environment, chance physiological phenomena can influence life span. Rea *et al.* report that the stress response system of *Caenorhabditis elegans* is subject to an underlying physiological randomness that affects how it copes with environmental insults. They placed the gene encoding green fluorescent protein (GFP) under the control of the regulatory region from the gene encoding a heat shock protein, creating an easily scored biomarker. Upon exposure to heat, isogenic worms exhibited considerable variation in fluorescence, and those expressing the highest amount of GFP tolerated heat

the best and lived the longest. The physiological state indexed by GFP expression level was not heritable, and the authors suggest that stochastic variation in molecular and biochemical reactions could account for the variation in individual robustness and longevity. — LDC

Nat. Genet. 10.1038/ng1608 (2005).

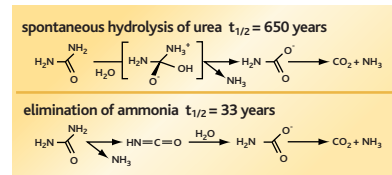
CLIMATE SCIENCE

Black Carbon

Global climate models are often used for detection and attribution experiments that assign cause to observed variations in climate. These studies have shown that most of the global warming that has occurred over the past 100 years has been caused by increasing concentrations of atmospheric greenhouse gases, and that warming has been moderated by the cooling effect of sulfate aerosols

(which reflect sunlight back into space). However, black carbon aerosols have not been included explicitly in these simulations, despite suspicions that they could have a significant effect on the global radiative energy balance, perhaps even outweighing that of sulfate aerosols, because black carbon, unlike sulfate, absorbs solar radiation and causes atmospheric heating.

Jones *et al.* report results from a detection and attribution analysis that includes black carbon aerosols, as well as sulfate aerosols and greenhouse gases. They find that black carbon is not as important as sulfate and that its inclusion does not change the conclusion that 20th-century warming is due mostly to the positive forcing of greenhouse gas variations. Nevertheless, the magnitude



The hydrolysis and elimination pathways of urea decomposition.

proficiency of the enzyme, calculated by dividing the biochemical quantity k_{cat}/K_m by the rate constant of the noncatalyzed reaction (k_{non}), is many orders of magnitude greater than that of the reigning

champion, ornithine 5'-monophosphate decarboxylase. In contrast, Callahan *et al.* have measured the hydrolysis of substituted ureas, which cannot undergo elimination, and extrapolated from these data to arrive at a much faster non-catalyzed hydrolysis rate—one that confirms urease as a proficient, but not record-setting, catalyst. — GJC

J. Am. Chem. Soc. 126, 6932; 10.1021/ja0525399 (2005).

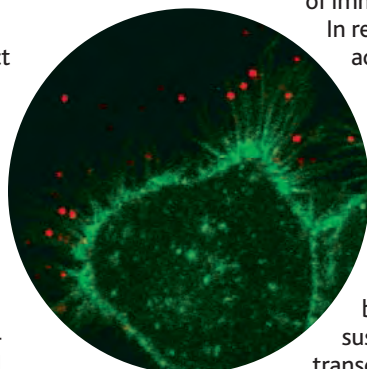
CELL BIOLOGY

Surf's Up

In order to infect a target cell, enveloped animal viruses must gain access to the cell's interior. The early stage of virus infection involves attachment to the cell surface and is frequently followed by endocytosis. Often, viruses are seen to bind to

cellular extensions such as microvilli or filopodia. Lehmann *et al.* asked whether such binding is a productive interaction for the virus, which needs to access the cell body (which can be far away) for successful infection. In vivo imaging studies revealed that after viruses bind to filopodia, they travel in a surfing type of movement along the cell surface toward the cell body, where they then can enter the cell. Filopodia are filled with actin microfilaments, and it is these filaments, in conjunction with cellular myosin II, that promote virus surfing. Disruption of surfing can reduce the efficiency of viral infection. — SMH

Fluorescently labeled viruses (red) attached to filopodia (green).



J. Cell Biol. 170, 317 (2005).

IMMUNOLOGY

A Matter of Choice

Thymocytes develop into two principal lineages: CD4⁺ or CD8⁺ T cells. In arriving at either fate, these cells first pass through a double-positive stage in which both CD4 and CD8 co-receptors are expressed, with one or the other later becoming permanently turned off.

To explain how this is regulated, Sarafova *et al.* extend their kinetic signaling model in which cell fate is determined by the context of T cell receptor (TCR) signals during the initial CD8 down-regulation

that takes place in all double-positive thymocytes. The model predicts that continued signaling in these cells (facilitated through TCR and CD4) would maintain CD4 transcription. However, if signaling were not sustained (as would be the case for TCR signals that depend on CD8 receptors), then CD4 expression would stop and CD8 transcription would resume. To test this, thymocytes from CD4-deficient mice were engineered to express a CD4 transgene under the control of immature CD8 transcriptional elements.

In response to CD4-dependent TCR activation, these cells down-regulated the CD4 transgene (as they also did for endogenous CD8), but subsequently re-started CD8 gene transcription to become functional CD8⁺ T cells. This supports the idea that regardless of TCR and co-receptor specificity, the fate of thymocytes is dictated by the presence or absence of a sustained T cell signal that mediates transcriptional cross-regulation of co-receptor expression. — SJS

Immunity 23, 75 (2005).

CHEMISTRY

Esters with Ease

Organic esters are widely used as fragrances and in the synthesis of pharmaceutical compounds. Among the many synthetic routes to esters, the oxidative dimerization of alcohols is direct and involves the endothermic liberation of a dihydrogen equivalent from each alcohol; efficient reactivity therefore requires another component, a stoichiometric H₂ acceptor.

Zhang *et al.* have developed a homogeneous ruthenium catalyst to couple primary alcohols into esters in the absence of any extra reagents. The reaction occurs in toluene (115°C) at 0.1 mole % catalyst loading; continuous purging of H₂ from the system drives the equilibrium to >90% yield of the ester for butanol, hexanol, and benzyl alcohol. The key component of the catalyst is a tridentate ligand, consisting of diethyl-amino and di-*tert*-butylphosphino coordinating groups appended to a pyridine ring, and the low kinetic barrier may be due to lability of the diethylamino arm at the Ru center. Preliminary studies of the mechanism support the initial oxidation of one alcohol to the aldehyde, followed by addition of the second alcohol to form a hemiacetal, which in turn loses H₂ to give the ester. — JSY

J. Am. Chem. Soc. 10.1021/ja052862b (2005).

Q Who's delivering science to every corner of the world?

Chris Bernau

Dr. Dinah Davidson



Steve Cook

“ Sharing one copy of *Science* around our research camp in Brunei requires a plan as systematic as the ants we're studying. On the boat, in a treetop, or on the deck after dinner, we all get our chance to catch up on what's new in science. ”

AAAS members Chris Bernau, Dr. Dinah Davidson, and Steve Cook

AAAS is committed to advancing science and giving a voice to scientists around the world. Helping our members stay abreast of their field is a key priority.

One way we do this is through *Science*, which features all the latest groundbreaking research, and keeps scientists connected wherever they happen to be.

To join the international family of science, go to www.aaas.org/join.



ADVANCING SCIENCE. SERVING SOCIETY

www.aaas.org/join

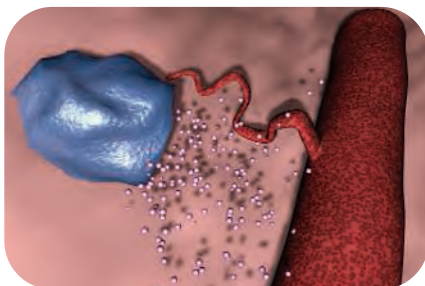
edited by Mitch Leslie

EDUCATION

Cells Gone Wild

Hungry tumor cells send out for dinner, releasing molecules that spur blood vessels to grow toward them (right). Learn more about the insidious ability—known as angiogenesis—and other aspects of cancer biology at this tutorial from lecturer Gregg Orloff of Emory University in Atlanta, Georgia, and contributors. The site is aimed mainly at cancer patients and health care workers but includes plenty of information for students. With an abundance of animations and diagrams, CancerQuest's 13 chapters plumb subjects such as the control of cell division and how defective genes bollix the delicate process. Visitors can also read up on clinical trials and experimental therapies, such as poisons that target only brain cancer cells carrying a particular surface receptor. Orloff is overhauling the site and will soon add new graphics and a timeline of cancer discoveries.

www.cancerquest.org



EDUCATION

Trapped Forever

Paleontologists prize amber because it serves as a sticky time capsule, entombing organisms from up to 300 million years ago. Archaeologists can use the fossilized tree resin to uncover ancient trade routes. You'll find a cache of amber lore at this site from librarian Susan Ward Aber of Emporia State University in Kansas. Amber forms when organic molecules oozed by trees react with oxygen and polymerize. The site covers topics such as where amber is found today—the Baltic area of Russia and the Dominican Republic are hot spots—and how to identify it. Real amber floats in saltwater, whereas plastic or glass imitations sink. Links at the Life in Amber section create a virtual gallery of animal and plant remains, from a 30-million-year-old grasshopper to the tiny flower of an extinct oak tree.

www.emporia.edu/earthsci/amber/amber.htm

RESOURCES

Britain's Birds

BirdFacts, a new guide from the British Trust for Ornithology, profiles 258 species that frequent or breed in the British Isles, such as the European coot (*Fulica atra*; right). The species accounts are crammed with ecological, anatomical, and conservation data. You'll find results from recent surveys of British and European populations and summaries of long-term trends in the species' numbers. The European coot, for instance, has been slowly increasing in Britain. Distribution maps compare censuses from the 1970s and 1990s and highlight range changes. Although it focuses on Britain, BirdFacts will prove useful for non-U.K. users because many of the species also inhabit Europe and North America.

www.bto.org/birdfacts



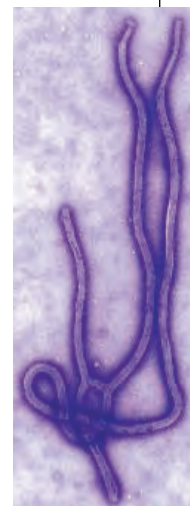
TOOLS

Deconstructing Viruses

If you're hunting for tools to analyze virus genomes and proteins, drop by The Viral Bioinformatics Research Center, created by Chris Upton of the University of Victoria in Canada. The site holds sequences for hundreds of viruses in 11 families, such as the Filoviridae, which includes the notorious Ebola virus (right). You can parse the data using 10 Java tools; for example, the

Base by Base program lets users compare viral genome sequences one nucleotide at a time. The site also offers background on the different families, describing their structures, life cycles, and how they hijack cellular activities. To learn more about some viral illnesses, download chapters from an infectious disease text.

athena.bioc.uvic.ca



DATABASE

Atlas of Other Worlds

They range from gas behemoths that dwarf Jupiter to a dainty body only about six times bulkier than Earth. Since the first one was detected a decade ago, the number of confirmed planets outside our solar system has climbed to more than 160, according to this database from Jean Schneider of the Paris Observatory in France. The Extrasolar Planets Encyclopaedia compiles vital statistics for each world, including mass and orbital axis, along with data for its parent star, such as spectral type and distance from Earth. The reports come from papers or preprints, conferences, and other planet-tallying sites. Visitors can also peruse a separate rundown of unconfirmed and retracted objects.

www.obspm.fr/planets

Send site suggestions to netwatch@aaas.org. Archive: www.sciencemag.org/netwatch

STEM CELL POLITICS

Frist's Support Raises Odds for Passage of Stem Cell Bill

Advocates of stem cell research are ecstatic about their new friend, Senator Bill Frist (R-TN). After failing all month to hold a promised vote on stem cell legislation, the Senate majority leader dropped a bombshell on 29 July, the last day before a 5-week recess: He announced that he supports expanding the number of embryonic stem (ES) cell lines eligible for federal funding. That puts Frist at odds with the policy laid down by President George W. Bush in August 2001 and on the side of those who hope the Senate will throw its weight behind a measure similar to one passed by the House of Representatives in May.

Scientists and even National Institutes of Health (NIH) officials have increasingly been chafing against Bush's policy as it has become clear that only 22 of the 78 cell lines he originally named are actually available for federally funded research. Many of these lines are aging and developing genetic abnormalities; all were derived with the aid of mouse feeder cells, which makes their clinical use problematic. A flurry of bills are vying for attention as senators who oppose the use of human embryos to derive stem cells offer alternative proposals to show that they're on the stem cell bandwagon. Now sponsors of S. 471, the Senate version of H.R. 810 that passed the House, are expressing confidence that the measure will sweep the Senate when it reconvenes in September—and possibly even garner the 67 votes necessary to override a presidential veto.

In a Senate speech, Frist said “the limitations put in place in 2001 will, over time, slow our ability to bring potential new treatments for certain diseases. Therefore, I believe the president's policy should be modified”—that is, federally funded scientists should be

allowed to work with ES cells derived after the presidential cutoff date of 9 August 2001. Frist added that ES cells “meet ... medical needs that simply cannot be met today by adult stem cells.”

His changed stance delighted stem cell supporters. “We had no idea he was going to make the switch,” said Representative Mike Castle (R-DE), a sponsor of the House bill (*Science*, 3 June, p. 1388). Senator Arlen Specter (R-PA), a dogged promoter of the Senate version, called Frist's action a political “earthquake.” Now others who follow his lead have “cover” and are far less likely to suffer political repercussions, Specter noted at a 29 July press conference.

In his remarks, Frist said that he's not

and Human Services (HHS) would issue after the bill is signed into law.

Scientists showered Frist with praise after his speech. “He is to be applauded,” says Stanford University stem cell researcher Irving Weissman, who says he had a phone conversation with Frist 2 days before the announcement in which the senator quizzed him extensively about stem cells. Weissman pushed hard on the need to allow research cloning, or somatic cell nuclear transfer experiments, before the two “agreed there were some areas where we'd always disagree.”

Much remains to be settled before S. 471 comes to a vote. Opponents of the bill have come up with at least six other measures that senators may be asked to vote on at the same time:

- S. 681, a bill already passed by the House, would authorize funding for a national cord blood stem cell network.

- S. 1557, introduced on 29 July by Senator Tom Coburn (R-OK), mirrors a House bill that would give NIH \$15 million to fund research on “alternative” ways of deriving pluripotent (ES-cell-like) cells.

- Senator Kay Bailey Hutchison (R-TX) plans to introduce a compromise of sorts between Bush's policy and S. 471. The measure would allow federally funded research on additional stem cell lines, but only those derived from frozen embryos currently held at fertility clinics.

- Senator Sam Brownback (R-KS) wants votes on two bills. One, S. 658, is a ban on all cloning, including research cloning—a measure that the House has already passed twice. The other bill (S. 1373) would ban the creation of what Brownback calls human-animal “chimeras.”

- Senator Dianne Feinstein (D-CA) on 27 July introduced a bill (S. 1520) aimed at banning just reproductive human cloning.

Frist has said he will introduce all the stem cell bills for up-or-down votes—that is, no amendments—but has not specified the timing. If Specter and Harkin don't get a clean vote on S. 471, they are prepared to play hardball by inserting the measure into the HHS 2006 appropriations bill, which Specter will be managing as chair of the relevant subcommittee.

—CONSTANCE HOLDEN

With reporting by Jocelyn Kaiser.



Brothers in arms. Frist (*inset*) is ready to support a bill backed by senators Orrin Hatch (R-UT) (*right*) and Harkin (*behind Hatch*).

entirely happy with H.R. 810. Although the bill is “fundamentally consistent” with principles he laid down 4 years ago, he said it lacks adequate ethical safeguards—in particular, the need to prohibit giving financial incentives to fertility clinics that could influence couples' decisions to donate “spare” embryos. A Frist spokesperson confirmed, however, that he is ready to support the bill in its present form. Senator Tom Harkin (D-IA) said that Frist's objections can be addressed in regulations that the Department of Health

864

Keeping
Alzheimer's
at bay

868

Mesopotamian
riddles

870

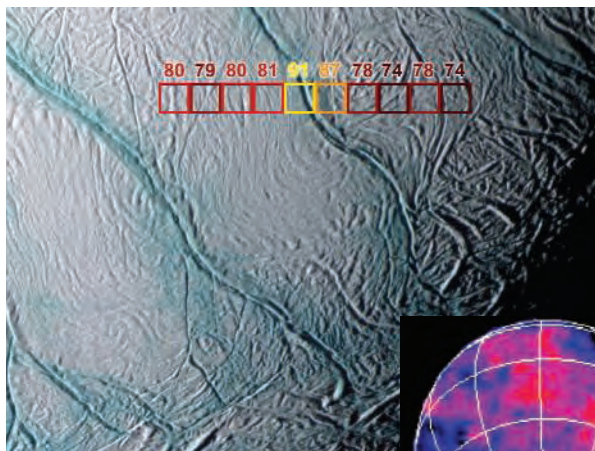
Can a flu
pandemic be
stopped?

PLANETARY SCIENCE

Cassini Catches Mysterious Hot Spot on Icy-Cold Enceladus

Since the Voyager flybys of the saturnian system in the early 1980s, planetary scientists have known there was something weird about the ice-covered moon Enceladus. Now, with Cassini's close flyby on 14 July, they understand the weirdness better. They can see places where the surface has recently been cracked and contorted. Some of those fractures near the south pole are so hot, by saturnian standards, that water vapor is spewing off the surface to form a tenuous atmosphere. However, the close-up encounter has only deepened the mystery of how a body as small as Enceladus can come up with enough energy for such an active geologic life.

Cassini unveiled the icy hot spot by bringing most of its instruments to bear on Enceladus during an exceptionally close 175-kilometer pass. The camera captured the most detailed images yet of the surface, revealing a network of cracks that look fresh enough to have formed in the geologically recent past.



Uncool moon. Warmth (yellow) appears in kilometer-wide fractures (above) near Enceladus's south pole (right).

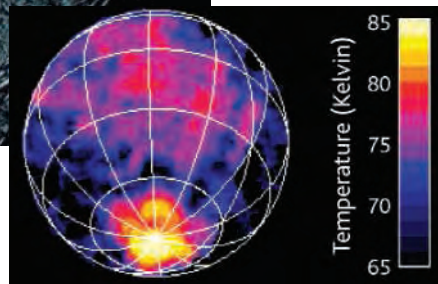
The infrared mapping spectrometer found a broad "hot" spot near the south pole that registered 85 K and reaches 110 K or above in places. "That's pretty spectacular for a body that should have temperatures of 60 K to 70 K," says Cassini science team member Torrence Johnson of the Jet Propulsion Laboratory in Pasadena, California. The infrared spectrom-

eter placed at least one tiny hot spot right on one of the cracks that the camera found near the south pole (see figure, left).

The hot spots of Enceladus are giving off more than heat. Three instruments independently detected an invisible plume of gas above the polar hot spot. The mass spectrometer detected water as Cassini passed the moon. The magnetometer found Saturn's magnetic field draped over an obstacle above the surface, pre-

sumably ionized water. And when a star passed behind Enceladus, the ultraviolet imaging spectrometer detected the absorption of its light by water. All three found the most water over the warm spot, not evenly distributed around the moon.

"The presumption is that we're looking at icy hot spots" near the south pole, says Johnson. Strictly speaking, Enceladus is not volcanically active, he notes. There's no sign of ice lavas having flowed across the surface or even of geysers spewing the water. But somehow cracks are being opened and heated enough ▶



PLANETARY SCIENCE

Newfound 'Tenth Planet' Puts Pluto Behind the Eight Ball

The discovery of a distant object larger than Pluto orbiting the sun seems secure enough. How to pigeonhole it, though, is completely up in the air. Is it the 10th planet, the first one discovered since Pluto in 1930? Or is it, with Pluto, just another Kuiper belt object (KBO), one of the thousands of icy chunks of debris left from the solar system's formation?

Three astronomers—Michael Brown of the California Institute of Technology in Pasadena, Chadwick Trujillo of the Gemini Observatory in Hilo, Hawaii, and David Rabinowitz of Yale University—first photographed the cosmic bone of contention almost 2 years ago. This January, they noticed that the object, temporarily designated 2003 UB313, was moving against background stars. They calculate that it is at the most distant point of its orbit—97 times as far from the sun as Earth is on average (97 astro-

nomical units, or AU)—and that its steeply inclined 650-year orbit will carry it as close as 36 AU from the sun. Pluto also follows an inclined orbit, between 30 and 50 AU.

The new object is so bright that it must be larger than 2390-kilometer-wide Pluto, the group reports. Because the orbiting infrared Spitzer Space Telescope cannot detect the new object, it must be smaller than 3200 kilometers.

Brown and colleagues are calling 2003 UB313 "the 10th planet." Some astronomers agree. "If it's larger than Pluto," says minor-planet astronomer David Tholen of the University of Hawaii, Manoa, "I'd call it the 10th planet, because Pluto is the ninth planet by historical precedent."

But theoretical astrophysicist Alan Boss of the Carnegie Institution of Washington's Department of Terrestrial Magnetism thinks

that would be a bad move. When the first asteroid, 946-kilometer Ceres, was discovered in 1801, he notes, astronomers called it a planet, too. But they demoted it to minor-planet status after other asteroids started showing up between Mars and Jupiter. The new object, Pluto, and several slightly smaller KBOs discovered recently "are all part of one population of objects," Boss argues—no one of which has enough mass and gravity to dominate its region of space the way the first eight planets do.

The International Astronomical Union in Paris, France, has been studying Pluto's status for 6 months, with no resolution in sight. Brown, who expects a few more larger-than-Pluto objects to turn up, is rooting for Pluto. "People love Pluto," he says. "Saying Pluto is not a planet will never be popular."

—RICHARD A. KERR

from below—perhaps by rising slush “magma”—that water molecules can escape to space from the relatively warm surface. Contrary to expectations (*Science*, 14 January, p. 202), this geologic activity seems to have nothing to do with creating the diffuse E ring of Saturn. (Cassini’s dust detector did find a cloud of dust around Enceladus, presumably chipped off the moon by micrometeorite

impacts, which might feed the E ring.)

Why Enceladus should be driving off its water from a south-pole hot spot remains a mystery. Neither of the usual sources of planetary heat—lingering decay of radioactive elements in deep rock, or tidal kneading by orbital interaction with Saturn and other moons—seems great enough, says planetary physicist David Stevenson of the California Institute of

Technology in Pasadena. Tidal heating is the more promising explanation, he says, but it heats other moons at least as much as Enceladus with no sign of a hot spot. So, appealing to tidal heating would require that “Enceladus is somehow special,” says Stevenson. “That’s uncomfortable.” Theoreticians will have to redouble their efforts to hammer out a moon they can live with.

—RICHARD A. KERR

AIDS RESEARCH

Male Circumcision Thwarts HIV Infection

A study in South Africa has shown for the first time that circumcising adult men can dramatically lower their risk of becoming infected by HIV through heterosexual sex. “It is a major advance in HIV-prevention studies,” said Catherine Hankins, an associate director at the Joint United Nations Programme on HIV/AIDS (UNAIDS).

For nearly 20 years, observational studies have suggested that circumcision protects men from HIV infection, but until now, there was no prospective evidence to support that conclusion. The new study, led by clinician Bertran Auvert of the University of Versailles in Saint-Quentin, France, began in August 2003 with more than 3000 uncircumcised men between 18 and 24 years old from the Orange Farm Township near Johannesburg. Half the participants were circumcised at the trial’s outset. As Auvert reported last week at an international AIDS meeting in Rio de Janeiro, the study was stopped early, in November 2004, after an interim analysis showed that “the protection effect of male circumcision was so high that it would have been unethical to continue.”

The study, performed in collaboration with Adrian Puren of South Africa’s National Institute for Communicable Diseases and funded by France’s Agence Nationale de Recherches sur le SIDA (ANRS), suggests that circumcision can offer 65% protection from infection. Only 18 men in the circumcised group acquired new HIV infections, as opposed to 51 in the uncircumcised group. Further bolstering the results, men in the circumcised group reported 18% more sexual contacts than controls. “It’s extremely exciting,” says King Holmes, an expert in sexually transmitted diseases at the University of Washington, Seattle. “It’s essentially an anatomic vaccine for life.”

Circumcision could profoundly curtail the spread of HIV in sub-Saharan Africa. Circumcision

practices vary greatly from country to country, and studies have shown that the regions with the worst AIDS epidemics have the lowest circumcision rates. But Auvert cautioned against rushing to integrate circumcision into public health policies. Not only are results from one location difficult to generalize, he said at the meeting, but the Orange Farm study also did not address whether circumcision reduces the ability of HIV-infected men to transmit the virus. That question is being explored in a Ugandan trial—one of three adult circumcision studies now under way.

Charles Gilks, who directs the HIV prevention and treatment program for the World Health Organization, emphasized another caveat. Adult circumcision carries serious risks, especially when performed by traditional healers who do not have proper training. And because circumcision does not provide complete protection, it could backfire if it encouraged men to have more unprotected sex—which would also greatly raise risks for

women. “We do need to make sure that we’re not hurting more than we’re helping,” said UNAIDS’s Hankins.

Even so, Helene Gayle, president of the International AIDS Society, which sponsored the conference, stressed that circumcision could be an important part of a comprehensive prevention strategy. “Obviously, there is no magic bullet,” says Gayle, who directs the HIV program at the Bill and Melinda Gates Foundation. “Prevention really is a combination approach where we need to put together all the things that we know can make a difference.”

Although the results from the new study have been eagerly awaited by AIDS researchers, *The Lancet* rejected a paper describing them for “reasons unrelated to the data and scientific content,” *The Wall Street Journal* reported last month. At issue, Auvert and Puren told *Science*, is an ethical disagreement that involves how participants learned their HIV status and the counseling they received. *The Lancet*’s rejection stunned the researchers. “We were taken aback,” says Puren.

The U.S. Public Health Service requires that when research it supports involves testing people for HIV, the participants must be informed of their results. Other funders, including ANRS, do not have this requirement. The Orange Farm study, which was approved by the ethics committee of South Africa’s Medical Research Council, provided counseling and also advised everyone to learn their status and receive more intensive counseling, but on a voluntary basis.

Auvert and Puren strongly defend the study. The researchers made these conditions voluntary because of concerns about the intense stigma that HIV-infected people often face in South Africa. “Many of these people prefer to be dead than rejected by their community,” says Auvert.

Lancet Editor Richard Horton declined to comment. But Ronald Gray, a reproductive epidemiologist at Johns Hopkins University in Baltimore, Maryland, thinks the journal reacted too harshly. “If there was an ethical lapse, I don’t think it was so egregious,” he says.

—JON COHEN

New HIV Infections				
Months	0–3	4–12	13–21	Total
INTERVENTION (circumcised)	2	7	9	18
CONTROL (uncircumcised)	9	15	27	51
TOTAL	11	22	36	69



First cut. Before circumcision becomes a prevention tool, the Orange Farm study’s positive results (above) must be confirmed by other trials now under way.

U.S. Government Shifts Stance on Claims to Ancient Remains

In an about-face, the U.S. government has sided with scientists fighting a proposal that would make it harder for them to investigate ancient human remains like those of Kennewick Man. At the same time, some scientific groups are supporting Native Americans in pushing for the proposal.

At a Senate hearing last week, Interior Department official Paul Hoffman spoke against a proposed broadening of the definition of "Native American" in the Native American Graves Protection and Repatriation Act (NAGPRA). The amendment, sponsored by Senator John McCain (R-AZ), would enable tribes to claim ancient human remains even if no genetic or cultural connections with living groups could be established. Native American groups began lobbying for the new definition after the 9th U.S. Circuit Court of Appeals in San Francisco, California, ruled in February that the 9400-year-old Kennewick bones, discovered in 1996, aren't covered by NAGPRA because they show no connection with any existing human group.

NAGPRA defines Native American as "of, or relating to, a tribe, people, or culture that is indigenous to the United States" (indigenous meaning pre-Columbus). Indians and their supporters want to add the words "or was" after "is." The seemingly small change would label bones such as those of Kennewick Man as "Native American" and might enable tribes to rebury them without allowing scientists to examine them. Last week's hearing came after scientists complained that the Senate Committee on Indian Affairs endorsed McCain's amendment, part of a package of changes to Indian-related laws (S. 536), without any public input.

The Bush Administration inherited the Kennewick case, in which the government, as the defendant in a case brought by a group of scientists, argued that the remains were covered by NAGPRA. But Hoffman indicated that the department, after 8 years of siding with the Indians, had changed its mind. "We believe that NAGPRA should protect the sensibilities of currently existing tribes ... while balancing the need to learn about past cultures," he testified. "[W]here remains are not significantly related to any

existing tribe, ... they should be available for ... scientific analysis."

Hoffman later told *Science* that "we thought [the appeals court] made a good argument." In addition to making the bones available to scientists, the appellate court ordered the NAGPRA grants program to pay the plaintiffs \$680,000 in attorneys' fees, Hoffman said.

Some scientists who testified at the hearing were sympathetic to the proposed word change. Keith W. Kintigh of the Society for American Archaeology in Washington, D.C.,



Mystery ancestor. This reconstruction of the 10,600-year-old, partially mummified body found in 1940 is the kind of discovery that could be lost to the rest of the world if it's handed over to an Indian tribe, say scientists.

said that requiring a demonstration of ties only to existing tribes is "inconsistent with a commonsense understanding" of the term Native American and might exclude "historically documented Indian tribes that have no present-day descendants." Patricia M. Lambert of the American Association of Physical Anthropologists agreed with Kintigh that everyone who inhabited America prior to the advent of Europeans should be treated as a "Native American."

In contrast, Portland, Oregon, lawyer Paula Barran, speaking for the Kennewick plaintiffs, cautioned that the proposed language would "stamp people from ancient cultures all as Native American as we know them." Any newly found human remains, she warned, would be "automatically turned over ... only to people calling themselves Native Americans." Such handovers have already been made under the current law, notes Douglas Owsley, a forensic anthropologist at the Smithsonian Institution, citing a set of 7800-year-old bones that the Minnesota Science Museum yielded up for reburial in 1994.

Owsley says that the vast majority of what are called "unidentifiable remains" are obviously Native American. But a few, like Kennewick Man, are what he calls "paleoamericans," far older than and racially distinct from Native Americans. "Many critically important skeletons will be forever lost if this bill becomes law," says Barran.

—CONSTANCE HOLDEN

ScienceScope

Bush Backs Teaching Intelligent Design

Echoing comments made about creationism when he was governor of Texas, U.S. President George W. Bush told a group of Texas reporters this week that he supports the teaching of intelligent design (ID) in schools. "Part of education is to expose people to different schools of thought," said Bush, without elaborating.

In February, John Marburger, the president's science adviser, stated that ID is "not a scientific theory." But Marburger says that he and Bush are not at odds over whether ID should be discussed in schools. "To ignore [ID] in the classroom is a mistake," said Marburger in an interview with *Science*, although he added that ID should not be taught "as an alternative" to evolution. Marburger wouldn't say whether he'd discussed the issue with Bush.

If Bush wants ID "to be a substitute or alternative [to evolution], ... that would be a terrible mistake," says Leonard Krishtalka, director of the Biodiversity Research Center at the University of Kansas in Lawrence. However, he notes, it's reasonable for the president to support teaching ID as part of the history of biology. ID proponent William Dembski of Baylor University in Waco, Texas, welcomed the support but said that he hoped Bush would support teaching ID in biology classes.

—JENNIFER COUZIN

Bird Flu Hits Russia

In a worrisome leap for the H5N1 avian influenza strain, Russian authorities have reported the first outbreak of the virus on their soil. The outbreak has killed thousands of chickens and wild birds around the Siberian capital of Novosibirsk; it appears to have started on 19 July at a lake in a village called Suzdalka where the two kinds of birds mingle. Russia's chief sanitary physician Gennady Onishchenko said at a press conference this week. A World Health Organization spokesperson says Russia should send samples from the outbreak to a lab outside the country to confirm the presence of the virus.

With many chickens in backyard pens, bird trading at markets, and poor infrastructure in rural Russia, it's unlikely that the country can contain the westward spread of the virus, which means it could reach Europe, says Ilaria Capua, a flu researcher at the Istituto Zooprofilattico Sperimentale della Venezia in Italy. "It's going to be very, very, very hard to stop it," she says.

—ANDREY ALLAKHVERDOV AND MARTIN ENSERINK

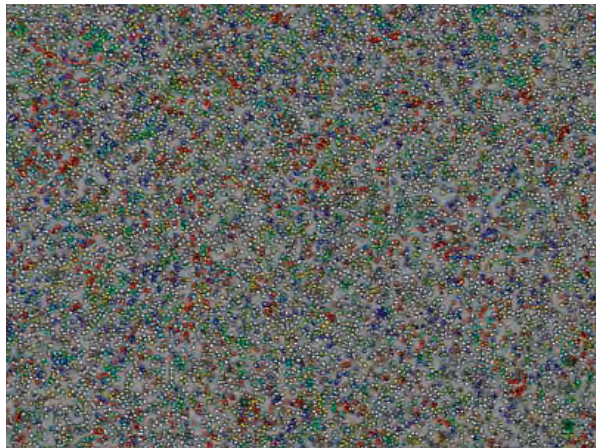
Cut-Rate Genomes on the Horizon?

Given the exorbitant cost of deciphering genomes, most labs have given up sequencing and left that job to the big sequencing centers. But now, two groups have published methods that may be much cheaper and faster, promising small labs a chance to do more of their own sequencing. "These are the first described techniques having the potential of replacing conventional [approaches]," says Mostafa Ronaghi, a biochemist at Stanford University in Palo Alto, California. And more are in the works, he adds.

In a paper published online by *Science* this week (www.sciencemag.org/cgi/content/abstract/1117389), George Church, a computational biochemical engineer at Harvard Medical School in Boston, and his colleagues introduce a do-it-yourself sequencer that uses a microscope and other off-the-shelf equipment. With this technology, his team sequenced a strain of *Escherichia coli* and was able to detect easy-to-miss single-base-pair changes from an almost identical *E. coli* genome. The approach reduces sequencing costs by 90%, Church says. Jonathan Rothberg, founder of 454 Life Sciences Corp. in Branford, Connecticut, has demonstrated the power of another cost-cutting technology he calls "454." He describes 454's success in sequencing *Mycoplasma genitalium* online in *Nature* this week.

Researchers all over the world still depend on a sequencing method introduced decades ago by Fred Sanger of the Laboratory of Molecular Biology in Cambridge, U.K. It uses bacteria to amplify the DNA and expensive reagents to label bases for identification. The cost has dropped since the mid-1990s from more than \$1 to less than a 10th of a cent per base. But it's still high for many projects, including medical uses such as checking the genomes of individuals.

Both groups save money by eliminating the need for bacteria and miniaturizing the process wherever possible. In lieu of bacteria, they attach DNA to aqueous beads encased in oil where chemical reactions copy the DNA to make the necessary amount. That change alone could reduce by two-thirds the costs associated with space and personnel, says Edward Rubin, director of the U.S. Department of Energy Joint Genome Institute in Walnut Creek, Califor-



Modern art, sequencing style. The color of each bead indicates the next identified base in a sequence. White beads have no DNA.

nia. Moreover, both perform many thousands of these sequencing reactions at once in miniature "reactors," decreasing the need for pricey chemicals.

Once the DNA is ready, the two technologies diverge: The 454 technique puts the beads on a fiber-optic chip and uses flashes of white light to identify the bases. Rothberg washes the chip surface with one base at a time, creating four light patterns that a computer puts together as a sequence. Church's technique uses bursts of different

fluorescent colors, one each to a particular base, to distinguish the bases. Both use high-speed charge-coupled device cameras to record the labeled bases.

Neither method is up to speed yet. The accuracy of both "should be improved by at least one order of magnitude," says Ronaghi. Also, to sequence mammalian genomes, the length of sequence generated, the "read," should be about 700 bases, but reads reported from these new approaches are hovering between 26 and 110 bases.

Rothberg's company has sold its \$500,000 machines to a dozen sequencing centers. Church's technology costs \$140,000 and is in use at three of those centers. However, says Rubin, there are still kinks in the build-it-yourself version. But, he adds, in time, it "may improve efficiencies and throughput even further [than the 454]."

Whatever their limitations, the two reports signal the dawn of a new era in genome sequencing and detecting changes in individual genomes. Last year, the U.S. National Human Genome Research Institute in Bethesda, Maryland, began a program aimed at decreasing the cost of sequencing mammalian genomes to \$100,000 in 5 years and to \$1000 5 years later. That's what many think it will take for sequencing to become affordable in small labs.

—ELIZABETH PENNISI

CLONING

The Perfect Pedigree

Afghan hound Snuppy (right) is the world's first canine clone, carrying the same DNA as his older twin Tai (left). Dogs have been difficult to clone, but Woo Suk Hwang of Seoul National University and his colleagues report in the 4 August issue of *Nature* that a new method of collecting oocytes—and persistence—paid off.



Hwang and his colleagues removed the DNA from more than 1000 canine oocytes, fused each with a skin cell from Tai's ear, and prompted the fused cells to begin dividing. They implanted 1095 resulting embryos into 123 surrogate mothers but detected only three pregnancies. One fetus miscarried, and two full-term puppies were born by caesarean section. The second dog died of aspiration pneumonia after being tube fed, Hwang says.

Although Hwang says Tai was chosen in part because of his "gentle and docile pedigree," he isn't interested in cloning friendly pets. Cloned dogs might help researchers find genes involved in hypertension or breast cancer, he says. And if scientists could grow embryonic stem cells from cloned canine embryos, the animals could serve as models for therapeutic cloning, in which genetically matched ES cells would be used to develop cells to replace those damaged by disease or injury.

—GRETCHEN VOGEL

U.S. Energy Bill Promises Some Boosts for Research

After a 4-year effort, Congress passed a landmark energy bill last week, setting out goals and incentives that could shape federal energy policy for the next decade.

The 1724-page bill includes \$14.6 billion in tax breaks—mostly to encourage domestic energy production from conventional sources—new efficiency standards for appliances, and renewed legal protections for nuclear power plant operators. It also contains provisions that aim to bolster federal spending on basic research, including an increase in the budget of the Department of Energy's (DOE's) Office of Science from \$3.6 billion in 2005 to \$5.3 billion by 2009; backs applied research efforts aimed at burning fossil fuel more cleanly; and calls for studies on combustion and carbon sequestration. But these commitments are far from assured: The bill simply authorizes spending that must later be approved by spending committees.

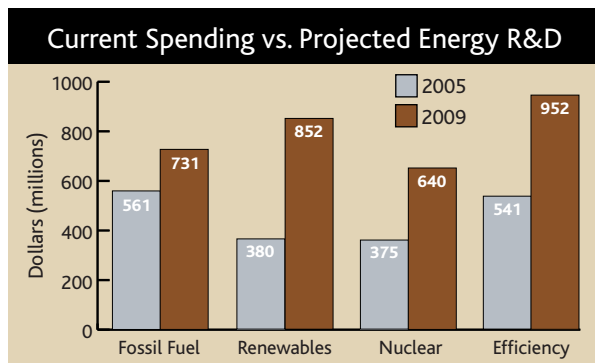
"The bill largely codifies the existing energy research programs," says a spokesperson for House Science Committee Chair Sherwood Boehlert (R-NY). While Boehlert helped craft the research provisions of the bill, he voted against it because it lacks government mandates that would boost energy production from renewables—such as wind and solar power—and raise fuel efficiency for automobiles.

For research, the bill lays out ambitious funding goals. If appropriations committees follow its lead—a big if—by 2009, fossil fuel research would rise by 23% and funding for renewables would more than double. The bill also sets a 2008 date for the construction of the Rare Isotope Accelerator, long sought by nuclear physicists to study exotic nuclei. Calling the R&D provisions "pretty good," William Fulkerson, former energy manager at Oak Ridge National Laboratory in Tennessee, says research projects on advanced coal plants and a plan to build a hydrogen-producing nuclear plant in Idaho should eventually help cut carbon emissions. President George W. Bush is expected to sign it shortly.

But critics say the bill fails to address basic questions about the direction of U.S. energy research. "The entire energy [bill] is really based on wishful thinking, that if you throw enough money at different technologies, one

of them will one day take over the market and solve our problems," says Joseph Romm, director of the Center for Energy and Climate Solutions in Arlington, Virginia. He argues that only mandates will force companies to develop better energy technology. New York University physicist Martin Hoffert, while praising the proposed increases, says the bill "lacks focus" and calls for a national debate on which energy research areas most deserve federal investment. Judy Biggert (R-IL), chair of the House Science Committee's energy subcommittee, defends the bill's multifaceted vision. "We can't rely on just one area of research," she says.

The bill's generous tax incentives represent a significant victory for the energy industries, although the final legislation did not keep protections sought by producers of gasoline additive methyl tertiary-butyl ether—contested provisions that



Promises, promises. The energy bill would give the biggest increase to renewables, but the money still has to be appropriated.

caused previous versions of the gargantuan legislation to fail. But utilities succeeded at keeping out a requirement that they draw 10% of their power from renewables. Jason Grumet of the National Commission on Energy Policy in Washington, D.C., says he fears that the only way to pay for the bill's hefty research increases is to generate income-producing measures such as auctioned carbon credits associated with a cap-and-trade system—a provision successfully blocked by energy firms.

Although environmentalists have railed against the bill, science lobbyists are cheering the creation of a new Under Secretary of Science position at DOE on the same level as the existing undersecretaries for defense and nondefense energy work, which should give researchers more influence in tough budget times.

—ELI KINTISCH

DeLay Hits Pay Dirt

Several science policy experts are criticizing a \$550 million program for oil studies created as part of the energy bill passed last week (see story on this page). Three-quarters of the funds in the 11-year program—whose research mission includes drilling, exploration, and other petroleum technology—are to be managed by a "corporation that is a consortium" chosen by the Department of Energy. But Representative Henry Waxman (D-CA) says the language is tailor-made for the Research Partnership to Secure Energy for America, an industry group located in the district of House Majority Leader Tom DeLay (R-TX), who supported the measure.

The Bush Administration has sought to cut oil and gas research, but the program will bypass congressional appropriators, drawing its funds directly from oil-lease income the government collects until the money runs out. Economist James Sweeney of Stanford University called the new program "pork" whose funding mechanism could set a harmful precedent for other energy research efforts.

—ELI KINTISCH

Beyond Kyoto

Last week in Vientiane, Laos, the United States signed a climate technology-sharing pact with China, India, and Australia. The agreement, which seeks to "complement the Kyoto Protocol," brings together for the first time major carbon emitters with top developing nations to share research from coal technology to geothermal energy.

The agreement will promote technologies to provide "clean, affordable, and secure" energy, the U.S. said in a statement, and Intergovernmental Panel on Climate Change chair Rajendra K. Pachauri welcomed the news. But critics attacked the agreement, also signed by Japan and South Korea, as toothless and undercutting Kyoto.

—PALLAVA BAGLA

Drop Tests, Congress Tells EPA

Congress has forbidden the Environmental Protection Agency from accepting any toxicity studies in which people were intentionally exposed to pesticides until EPA issues a final rule spelling out its policies. The legislators said draft rules released by EPA last month (*Science*, 8 July, p. 232) didn't offer enough protection to research subjects. The restriction was part of a spending bill finalized last week.

—ERIK STOKSTAD

Recent research suggests that keeping mentally and physically active when young and middle-aged can help stave off the brain degeneration of Alzheimer's

Preventing Alzheimer's: A Lifelong Commitment?

A sweating man feverishly pumping an exercise bicycle may not seem to have much in common with a chess player coolly contemplating her next move. Yet both may be protecting their brains from the ravages of Alzheimer's disease. Recent results, some from epidemiological studies and others from investigations of animal models of Alzheimer's disease, suggest that exercise—both physical and mental—can help the brain combat the pathological changes that cause the illness.

If so, then people who engage in physical exercise and intellectual activities such as reading, solving crossword puzzles, and playing cards or chess may be able to slow down the development of Alzheimer's disease, perhaps delaying it long enough that incapacitating symptoms won't appear during a person's lifetime. "The brain is an organ that, like any other organ, ages depending on how it's used," says neurologist Robert Friedland of Case Western Reserve University School of Medicine in Cleveland, Ohio.

Yet parts of the story may not be that simple. Researchers are debating, for example, whether intellectual activities are actually protective or whether people who participate in them are more resistant to Alzheimer's disease, possibly because of the way their brains developed.

Building a cognitive reserve

Parents who warn their children that they will regret not going to college could be correct—but in an unexpected way. Over the years, several studies have shown that formal education seems to protect against Alzheimer's disease. For example, a 1997 study of 642 elderly people, conducted by Denis Evans of Rush Presbyterian—St. Luke's Medical Center in Chicago and his colleagues, found that each year of education reduces a person's risk of Alzheimer's disease by 17%.

As suggested in the late 1980s by Robert Katzman of the University of California, San Diego (UCSD), education might protect against Alzheimer's disease by increasing the number and strength of neuronal connections in the brain, thus improving an individual's

so-called cognitive reserve. According to this theory, later in life when Alzheimer's pathology begins to eat away at the brain's neurons, people with larger reserves would be better able to cope with the onslaught.

One recent study supporting Katzman's idea came 2 years ago from a Rush Presbyterian—



Healthy bodies, healthy minds? Some studies show that exercise can slow cognitive decline.

St. Luke's team led by David Bennett and Robert Wilson. Since the mid-1990s, these researchers have been following a group of older Catholic priests, nuns, and brothers who had agreed to donate their brains after they died.

Analysis of the brains available in 2003, 130 in all, showed no correlation between education and the formation of plaques and tangles, the abnormal brain deposits that characterize Alzheimer's disease. But a battery of 19 tests performed periodically in the years before the donors died revealed that people with high levels of education better maintained their cognitive abilities. Wilson says that the highly educated participants

didn't develop Alzheimer's disease until they had about five times as many plaques and tangles as the less educated participants. "This suggests that education or cognitive activities achieve their effects by helping the brain tolerate the pathology," he says.

Not everyone finds support for the cognitive reserve theory, however. The so-called Nun Study points to a different conclusion: Early variations in how the brain develops makes some brains more resistant to developing Alzheimer's pathology than others.

David Snowdon started the Nun Study more than 15 years ago, when he was at the University of Minnesota, St. Paul. It originally included 678 members of the School Sisters of Notre Dame, all of whom were born before 1917. Snowdon, now at the University of Kentucky in Lexington, reasoned that studying nuns would help him identify factors that influence Alzheimer's development because they all have similar lifestyles and medical care. This eliminates some variables, such as smoking, that might skew the results.

As in other studies, Snowdon and his colleagues found that high education levels seem to protect against Alzheimer's disease. The researchers originally thought that this supported the idea that more education leads to a higher cognitive reserve. But analysis of biographical essays the sisters had written when they entered the convent, usually in their early 20s, pointed in a different direction. The early writings, Snowdon says, were an even better predictor of who would get Alzheimer's disease than education level. "Those who had the lowest linguistic skills at age 22 had a very high risk of Alzheimer's," Snowdon says. Indeed, most of the cases occurred in the nuns whose essays put them in the bottom third on the linguistic ability scale.

When Snowdon, neuropathologist William Markesbery, also at the University of Kentucky, and their colleagues examined the brains of nuns who had died, they found that those of lower linguistic ability were also much more likely to have signs of Alzheimer's disease such as brain shrinkage and tangles, although not plaques. That finding

CREDITS (TOP TO BOTTOM): A. PAKIEVA/PHOTO RESEARCHERS INC.; JOHN HENLEY/CORBIS

took Snowden by surprise. He points out that if the lower risk of Alzheimer's disease in the high-linguistic-ability group was solely due to their having a better cognitive reserve, the pathology ought to be similar in all the nuns. Instead, it appeared as if the brains of the sisters with higher linguistic ability were somehow more resistant to developing the pathology in the first place. This, Snowden suggests, might reflect differences in how the brain develops before birth and in early life. "Ultimately, it gets down to brain wiring and the biological mechanisms that defend the brain from disease," he says.

That doesn't mean, however, that there's nothing we can do to decrease our likelihood of getting Alzheimer's disease. "Genes are the driving force, but it's highly likely that diet and lifestyle influence Alzheimer's risk," Snowden says. One indication of this comes from Margaret Gatz of the University of Southern California in Los Angeles.

In work she described at the International Conference on Prevention of Dementia held 2 months ago in Washington, D.C., she and her colleagues identified 109 pairs of identical twins in the Swedish Twin Registry in which one had been diagnosed with dementia and the other had not. "We do find that there is a difference in education. The twin with dementia had significantly less," Gatz says. So even in these genetically identical individuals, education apparently pays off in lowered Alzheimer's risk.

Several additional studies by teams including Friedland's, the Rush Presbyterian group, and Herman Buschke and his colleagues at the Albert Einstein College of Medicine in New York City suggest that a lifelong commitment to intellectual activities may aid in—indeed, may even be necessary for—maintaining any protection against Alzheimer's disease accrued in early life. "All forms of leisure activities requiring mental activity—reading, puzzles, cards, board games, crafts—are protective," Friedland says. "I believe they all involve learning in some way."

Conversely, Friedland and his colleagues found that one leisure activity that is arguably not intellectually demanding—watching television—was associated with an increased likelihood of developing Alzheimer's disease. Using questionnaires, they surveyed 331 normal controls and also the close associates, primarily spouses and children, of 135 Alzheimer's patients to find out what activities they participated in during midlife. As reported in the July issue of *Brain and Cognition*, the patients had watched more television; each additional hour of watching per day increased the Alzheimer's risk by a factor of 1.3. That doesn't necessarily mean that heavy television watching rots the brain. Rather, Fried-

land says, it may be a marker for an intellectually inactive lifestyle.

These epidemiological studies all suffer from the same complicating factor, however: Much evidence—including the Nun Study and a meta-analysis of 47 studies reported in the 31 July issue of *Neuropsychology* by a team led by Lars Bäckman of the Karolinska

Since then, additional studies have borne out that conclusion. These include two large prospective epidemiological studies that focused on women. In one, Kristine Yaffe and her colleagues at UC San Francisco, followed for 6 to 8 years nearly 6000 women over age 65 who did not show signs of Alzheimer's disease at the time they were



Mind matters. A study of nuns suggests that high linguistic ability early in life correlates with lower Alzheimer's risk; engaging in lifelong mental activities, including crafts, may also help stave off the disease.

Institute in Stockholm, Sweden—indicates that Alzheimer's disease develops slowly over many years before failing memory and other symptoms become apparent. Although researchers have conducted long-term prospective studies that try to exclude people already showing Alzheimer's symptoms, it is hard to eliminate the possibility that low participation in cognitively demanding activities may be an early symptom of the disease rather than a cause.

More reason to exercise

Pursuing an intellectual life may not be the only tack that people can take to ward off Alzheimer's disease. Some recent research indicates that physical exercise can be as good for the mind as for the body, although the literature on this issue has been mixed, with not every study showing a benefit.

A few years ago, Arthur Kramer of the University of Illinois, Urbana-Champaign (UIUC), and his colleagues performed a meta-analysis of 18 trials involving adults between the ages of 55 and 80 that explored the effects of physical exercise on performance of various cognitive tasks. They concluded that the answer to the question, "Does aerobic exercise enhance cognition?" was an "unequivocal yes."

recruited into the trial. The other trial comes from Francine Grodstein of the Harvard School of Public Health and her colleagues, whose study group included 18,766 women aged 70 to 81 from the Nurses' Health Study. Both studies reached the same conclusion: Women who got the most exercise, mainly walking, showed less cognitive decline over the years than women at the low end of the activity scale.

Varying one's exercise routine may also have mental benefits beyond relieving boredom. Constantine Lyketsos of the Johns Hopkins Medical Institutions in Baltimore, Maryland, and his colleagues have looked at the effects of physical activity on mental abilities in more than 3000 men and women in the Cardiovascular Health Cognition Study. "What mattered wasn't the absolute energy expenditure but the number of activities," Lyketsos says.

As reported in the April issue of the *American Journal of Epidemiology*, study members who engaged in four or more physical activities, which could be anything from gardening to jogging or biking, had about half the risk of dementia as that of participants who engaged in one or none. The effect was primarily seen, however, in persons who did not carry a gene variant called *ApoE4* that's

known to increase Alzheimer's risk. In the *ApoE4*-endowed population at least, genetics seems to trump activity.

The exercise studies all have the same potential downfall as the ones focusing on education and mental activities: the possibility that low activity levels are an early sign of Alzheimer's disease rather than a cause. But the exercise conclusions receive additional support both from imaging studies of human brains and from investigation of animal models of Alzheimer's disease.

Neurobiologists have known for some time that the human brain shrinks with age. Between ages 30 and 90, the losses range from 15% to 25% of brain matter, with the shrinkage particularly severe in areas such as the frontal and temporal cortex that are involved in memory and learning. About 2 years ago, Kramer and his colleagues confirmed such cortical shrinkage by using magnetic resonance imaging (MRI) to observe the brains of 55 older adults. But they also found that those losses were much reduced in the most physically fit individuals (assessed by performance on a treadmill).

In a second set of experiments, Kramer and his colleagues used functional MRI to assess brain activity in subjects performing a cognitive task. The more physically fit individuals not only performed better on the task than the less fit participants, but their brains also showed higher activity in the areas associated with the task. "Fitness training improves neuronal efficiency and performance," Kramer says. "Older brains are a lot more flexible and plastic than we have been led to believe."

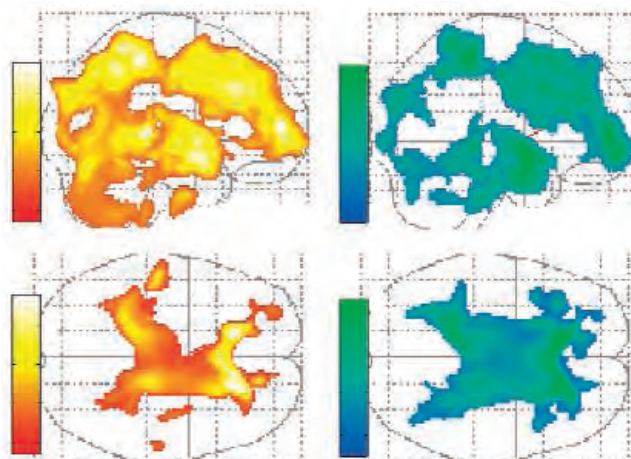
Lessons from rodents

How physical exercise enhances brain maintenance and function is unclear, but work with rodents points to several possibilities. One is that it improves cerebral blood flow, thus providing better nourishment to the neurons. For example, William Greenough and colleagues at UIUC have shown that exercise increases blood flow to rat brains, at least partly by stimulating the growth of the capillary vessels feeding the cortex.

Another possibility is that exercise turns up production of proteins that stimulate neuronal growth. About 10 years ago, Carl Cotman's team at UC Irvine, found that the brains of rats who ran voluntarily on a wheel show increases in one such factor, BDNF (for brain-derived neurotrophic factor). The increase was particularly strong in the hippocampus, an area involved in

learning and memory that's hard-hit by Alzheimer's disease.

Consistent with that finding, researchers including Greenough and Fred Gage at UCSD, have shown that exercise or so-called



Brain preservation. The brain images at left show the areas of gray matter (top) and white matter (bottom) that shrink with age. As indicated by the images at right, cardiovascular fitness can help preserve those brain regions.

enriched environments in which animals live in cages equipped with exercise wheels and other toys can increase formation of brain neurons and lead to other changes that should strengthen neuronal connections.

For example, the numbers of dendrites, the tiny projections of nerve cells that receive incoming signals, normally decline with age. But Greenough's team found that keeping rats in an enriched environment could counteract that decline. All but the oldest animals kept in such an environment showed increases in dendrite numbers, and even those very old rats maintained their dendrites better than control rats kept in standard lab cages equipped with nothing more than food, water, and bedding. "At the younger ages, use it [the brain] and gain," Greenough says. "And at the older ages, use it or lose it."

Exercise may even prevent formation of the amyloid typical of Alzheimer's disease, although the evidence, derived from animal models, is far from clear. Sam Sisodia's team at the University of Chicago in Illinois has been studying mice genetically modified to overproduce a protein called β amyloid, a major plaque constituent. As the researchers report in the 11 March issue of *Cell*, mice kept in an enriched environment produced much less of the protein and had fewer plaques than did animals kept in standard cages (also see *Science*, 11 March, p. 1547). The effect was especially pronounced in those animals who spent the most time on the running wheels, Sisodia says.

The Sisodia team didn't determine whether the enriched environment improved the animals' learning abilities, but in a similar

experiment, Cotman and his colleagues did. The Irvine group found that voluntary exercise, again running on a wheel, not only decreased the number of plaques in the hippocampus and cortex of mouse brains but also improved a rodent's performance on a cognitive task, learning to find a hidden platform in a water maze. (The results appeared in the 4 May *Journal of Neuroscience*.)

But not everyone has found that an enriched environment results in decreased plaque formation in Alzheimer's mice. In 2003, Joanna Jankowsky of the California Institute of Technology in Pasadena, David Borchelt of Johns Hopkins University School of Medicine in Baltimore, and their colleagues reported just the opposite: that it can lead to increased plaque formation. That experiment has been criticized on the basis that the animals were under stress. Jankowsky disputes that, noting that her team has since done

additional experiments with a different genetically altered mouse strain. Again, they found that β -amyloid and plaque deposition increased when compared to control animals kept in standard cages. "Not only have we found the same result, but we found it in another strain of mice," Jankowsky says.

Even so, the experiment sounds a hopeful note about the effects of enrichment. All the animals kept in the enriched environment showed improved performance on three different cognitive tests, although the mice with high β -amyloid production fared less well than animals with lower levels.

The reason for the discrepancy between the various groups' plaque findings are unclear, although it might be due to differences in the strains of mice used. Still, Cotman describes the recent results as "cool." He points out that, taken together, they indicate that it may be possible to prevent or slow the mental decline of Alzheimer's disease with or without decreases in plaque formation.

Researchers warn, though, that people need to act before they get old. "If you're going to do something to ward off Alzheimer's, you have to do it before memory problems develop," Snowden says. On the bright side, the interventions to prevent Alzheimer's disease are looking pretty much like the same ones recommended to prevent obesity and cardiovascular disease. Yaffe, who runs a memory clinic, notes that she tells her patients that exercise is "inexpensive, has very few side effects, and if worst comes to worst, it's good for your body."

—JEAN MARX

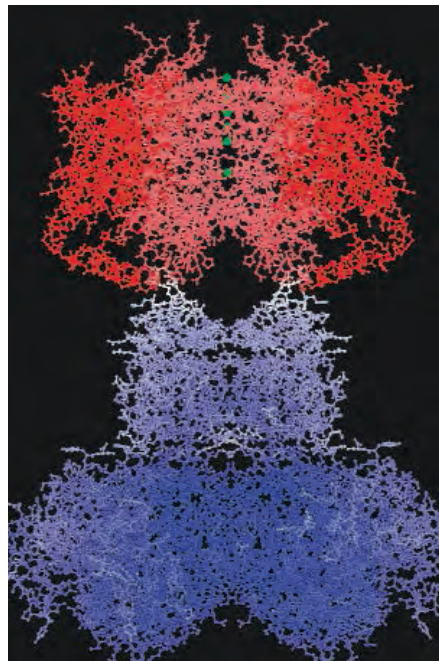
A New Portrait Puts Potassium Pore in a Fresh Light

An atomic structure of a potassium channel, which is central to the firing of neurons, may settle some debates over how this critical protein operates

Two years ago, Roderick MacKinnon and colleagues at Rockefeller University in New York City unveiled an atomic map of an ion-channel protein that rocked the small community of researchers working to unravel the details of how nerves and other cells conduct electrical impulses. Solving that structure was a tour de force of chemistry. It required getting copies of the protein, which is normally embedded in the bacterial cell membrane, to arrange in an ordered crystal. But the map showed parts of the molecule to be in positions that were not what other experts had come to expect. Now, in a pair of papers on pages 897 and 903, MacKinnon and colleagues reveal the structure of a closely related eukaryotic ion channel protein that has other experts breathing much easier. “This is a terrific structure,” says Richard Horn, a physiologist at Jefferson Medical College in Philadelphia, Pennsylvania.

Both the new and earlier structures offer close-ups of proteins called potassium voltage-gated ion channels. When a neuron fires, potassium ions build up inside the cell. These positively charged ions create an electrical voltage across the cell membrane that triggers the potassium channel to open, allowing the excess potassium ions to spill out and restoring the cell to its resting state, ready to fire again. Virtually all potassium channels consist of four identical subunits, each made up of six linked helical segments. Two of those helical segments from each subunit assemble to form the central pore through which ions flow, while the other segments form a cloverleaf pattern of four voltage sensors that detect voltage changes across the cell membrane and move like a lever to open and close the pore.

Getting structures of such complex proteins is no simple task. Researchers must first coax billions of copies of a protein to stack in a perfectly ordered crystal. They then fire a tight beam of x-rays at the crystal and track how those x-rays ricochet off the atoms in the crystal to work out the precise positions of each atom. That task is especially challenging with potassium ion channels, MacKinnon says, as the voltage sensors are barely connected to the pore regions and therefore are floppy and difficult to stabilize in a crystal. Two years ago, MacKinnon’s group attached antibody fragments to copies of the bacterial protein to



Hot shot. Potassium channel (red) is seen together with T1 and β domains (blue) for the first time.

help stabilize it. But when the structure was published in *Nature* in 2003, the voltage sensors looked to many experts to be tilted on their sides from what they expected (*Science*, 27 June 2003, p. 2020). Many of the measurements that had been done on the channels over the years didn’t seem to mesh with the new structure, Horn says. MacKinnon and colleagues themselves noted in the *Nature* paper that portions of the protein were in unexpected positions, possibly as a result of the technique used to crystallize the protein.

For their new structure, MacKinnon’s team was able to do away with the antibodies. The eukaryotic channels are nearly identical to those in bacteria, but there’s a key difference: Eukaryotic potassium channels contain an additional protein domain, known as T1, and another associated protein, known as β , that sit outside the cell membrane in the cytoplasm. With the help of some novel crystallization techniques that used lipids to crystallize the entire complex, MacKinnon’s team found that T1 and β helped stabilize the channel protein during crystallization without requiring support from antibodies. Whether

or not getting rid of the antibody fragments made the difference, the voltage sensors in the new structure are rotated upright, where other lines of evidence suggested they should be. “It was comforting to see [the position of the voltage sensors] was much more like everyone thought,” says Francisco Bezanilla, an ion channel researcher at the University of California, Los Angeles (UCLA).

Like its predecessor, the new structure offers fresh insights into how the channel works. For one, Horn says, helices that form each subunit’s voltage sensor aren’t adjacent to those that help make up the pore. Rather, those domains from the different subunits interlace around one another. Gary Yellen, a neuroscientist at Harvard University adds that the new structure shows for the first time how the voltage sensor links to the pore, which, he says, “is a pretty neat thing to see.”

Controversies remain. For example, ion-channel experts have long known that four positively charged arginine amino acids sit atop each of the voltage sensors that surround the pore. These charged arginines move in response to changes in the voltage across the cell membrane, pressing up and down on the lever that opens and closes the pore. But just how this movement takes place remains at issue.

MacKinnon’s team has suggested that two of the helices that help make up the voltage sensor are part of a “paddle” that moves through the membrane and pushes on the lever. That view, he suggests, was supported by a study 2 years ago that showed that positively charged arginine amino acids that are part of this paddle move a considerable distance—15 or so angstroms—through the membrane, from the extracellular to the intracellular portion, as the pore moves from its open to its closed configuration. David Clapham of Harvard University says the new structure is consistent with this model. But not everyone is convinced. Bezanilla points out that two 1999 studies, by his group and Ehud Isacoff’s group at UC Berkeley, used fluorescence tracking techniques to show that a key helix in the sensor that presses on the lever, known as S4, does not change its depth in the membrane by more than 3 angstroms.

MacKinnon points out that this debate can’t be settled by the new crystal structure because it’s a static view of the potassium channel in the open position. He says he and his colleagues are already working to get a structure for the protein in its closed form, which together with the current structure should reveal how the protein moves. In solving this debate, MacKinnon says, “there is nothing like data.”

—ROBERT F. SERVICE

Alas, Babylon: Tracing the Last King's Desert Exile

Mid-6th century B.C.E. was a dark time for the empire of Babylonia. Persians and Medes were threatening in the east, and the king mysteriously abandoned his famed capital of Babylon for a remote oasis in the western

60 centimeters (cm) wide, 50 cm high, and 11 cm thick—was later reused in building a wall. Only about a dozen lines of the stele are legible, but they indicate that Nabonidus made offerings to Babylonian deities—

including Marduk—in the form of carnelian, lapis lazuli, and censers of gold, according to a translation by Assyriologist Hanspeter Schaudig of the University of Heidelberg in Germany. The find “is very valuable for our knowledge of history,” says philologist David Weisberg of Hebrew Union College in Cincinnati, Ohio. But he adds that the inscription “is quite damaged, and many lines are illegible,” so it will require more study.

The find is part of a larger effort to understand the complex trade routes that linked the ancient Middle East. Tayma lies at a critical juncture of the frankincense trade flowing north from Yemen and other routes to the

Persian Gulf and Mesopotamia, and for millennia it offered travelers a respite from the desert. At the time of Nabonidus, the oasis included a city with a vast wall some 14 kilometers in circumference and a well 18 meters across, one of the largest on the notoriously dry Arabian Peninsula. The team, led by Ricardo Eichmann of Berlin's German Archaeological Institute and Said al-Said, a professor at King Fahd University, has found 13 successive layers of occupation from the mid-3rd millennium to the early centuries of the modern era, showing a surprising continuity in urban desert life.

Although Babylonian texts mention that Nabonidus built a palace at the site, Eichmann says none has yet been found, but the team will keep looking when it returns to Saudi Arabia in November. Textual evidence found elsewhere indicates that Nabonidus was ill when he left Babylon and recovered during his decade in the desert. But German excavation director Arnulf Hausleiter speculates that his real motives could have been economic: By asserting control over an important trade city, Nabonidus may have been attempting to bolster Babylon's flagging treasury. If so, the gambit failed. The texts say that the king returned to Babylon in

CHICAGO, ILLINOIS—More than 300 Mesopotamian scholars gathered at the University of Chicago's Oriental Institute from 17 to 23 July.

542 B.C.E. after a decade in exile, only to be overthrown by the Persian King Cyrus the Great 3 years later. Thus Mesopotamians lost control over their own rich territory—a control that was not fully regained until 2500 years later in the 20th century.

Ur's Xena: A Warrior Princess for Sumeria?

One of the most spectacular archaeological discoveries in history was Leonard Woolley's excavation of the royal tombs of Ur in the late 1920s. The 16 graves included a “death pit” with sacrificed retainers and animals. Woolley believed the tombs were those of kings and their consorts, including the famous Queen

Puabi, buried with a magnificent crown and other jewelry.

But one grave, tomb 1054, left Woolley perplexed. In the shaft 4 meters above the stone burial chamber was a cylinder seal inscribed with the word “lugal,” Sumerian for “king” or “ruler,” along with a name read as Meskalamdug and traditionally translated as “hero of the land.” In the stone chamber itself were a host of weapons, including a dagger at the side of the principal occupant. But there was one hitch: Woolley determined that the remains were of a woman. Scholars had long held that ancient Mesopotamian rulers, unlike their Egyptian neighbors, were always men. “That seal cannot be hers,” Woolley concluded in a 1934 publication.

The puzzle has obsessed two generations of researchers, who have come up with a variety of theories to explain it. Now Kathleen McCaffrey, a graduate student at the University of Cal-

ifornia, Berkeley, says that the most logical answer is the simplest: The seal and weapons did

Fit for a princess? The mysterious occupant of tomb 1054 wore this gold dagger at her side.



King's record. Ricardo Eichmann studies the stele that records Nabonidus's exile.

Arabian desert. Contemporary texts portray King Nabonidus as mentally unstable and complain that he forsook the prime Babylonian deity, virile Marduk, for the mystical cult of the moon god Sin, often portrayed as an old man with a long beard.

Those texts, written by Nabonidus's clerical enemies, have been the only evidence of his claimed exile. Now archaeologists have found the first concrete signs that Nabonidus indeed lived in the oasis of Tayma, more than 1000 kilometers to the west of today's Iraq, and they hope also to uncover why this obscure oasis played such a pivotal role in history. Academics familiar with the Middle East say that the Tayma dig itself, in sparsely settled northwestern Saudi Arabia, is a triumph of science over politics, given the difficulty of winning permits from the Saudi government for excavations by foreign teams.

Three years ago, Saudi researchers working near Tayma found rock inscriptions that mention an army of Nabonidus that battled local Bedouin. Then in December, a joint Saudi-German team found a piece of badly weathered stele, a stone slab inscribed with writing, which closely resembles other slabs associated with Nabonidus's reign.

The slab originally would have stood for passersby to read, but the team's fragment—



CREDITS (TOP TO BOTTOM): DAI ORIENT-ABTEILUNG; K. MCCAFFREY

indeed belong to the buried skeleton, which may have been that of a female Sumerian ruler. That claim has sparked fierce debate, however, especially because Woolley disposed of the bones shortly after discovering them.

Woolley himself suggested that the seal and weapons were gifts from the woman's husband. Another theory is that the true owner of the seal, a male, was buried in a mud-brick shaft above the stone tomb. But McCaffrey notes that the materials in that shaft are low quality and lack weapons, and that no other royal tomb is constructed of mud brick. In fact, the remains in the mud-brick shaft, identified by Woolley as male, were wrapped in women's clothing with feminine jewelry. Unfortunately, those bones also were discarded.

The principal occupant of 1054 herself reveals some curious gender anomalies, notes McCaffrey. Her skeleton was found wearing a hair ribbon, two golden wreaths, and a gold dress pin, all typical for high-status Sumerian women of the day. But she was not adorned with the usual earrings or elaborate choker, and there were no floral combs or cosmetic containers. And a gold headpiece and a dagger and whetstone at her waist were typical for Sumerian men; a gold headdress near the skeleton has a brim, a style that Woolley believed was worn mostly by men.

Also in the stone chamber were a bronze ax, dagger, and hatchet—very atypical for a woman's tomb. Other researchers attribute those weapons to the male attendants in the room, but McCaffrey notes that the attendants lack rings, weapons on their bodies, or any other sign of elite materials, suggesting that they were servants.

McCaffrey maintains that the root of the problem is translation: Sumerian grammar does not include gender distinctions, but “lugal” has always been translated as “king” rather than simply “ruler.” In the case of tomb 1054, she concludes that the woman was in fact a lugal.

But other scholars hotly disagree. University of Chicago archaeologist McGuire Gibson argues that the seal's location above the stone chamber makes it difficult to tie it to the elite occupant below. He adds that most of the bones had deteriorated so much that identifying gender was difficult. “Woolley couldn't tell the difference between a man, a woman, or a monkey,” he says. McCaffrey counters that Woolley was competent enough to identify correctly the genders of the dozen skeletons that still exist. Philologists, meanwhile, note that although “lugal” is technically a gender-free term, there is the counterpart term “eresh,” which traditionally is translated as female consort to a male ruler.

Without a skeleton, scholars may never definitively sort out the mysteries of tomb 1054. But the women of ancient Ur may have more to say in the near future: Researchers are now examining Queen Puabi's remains for clues to her genetic identity.

Looted Tablets Pose Scholar's Dilemma

Few societies before our own were as obsessed with recording data as ancient Mesopotamia. After inventing the first script in the 4th millennium B.C.E., the Sumerian scribes used clay tablets to keep track of the most minute economic transactions as well as great myths such as The Epic of Gilgamesh that stir readers even



Stolen. Looted cuneiform tablets, like these recovered in Jordan, are pouring out of Iraq.

today. The tablets have proved invaluable in understanding the hearts and minds of that lost world.

But the artifacts also have attracted collectors and antiquities dealers. Today, as many as 100,000 tablets a year are being ripped out of archaeological sites in war-torn Iraq and put on the international market, according to U.S. government estimates. By comparison, only some 300,000 to 400,000 likely existed in libraries and private collections prior to 1990, say scholars. So far, the number of stolen tablets confiscated or returned is minuscule: An FBI official said at the conference that fewer than 400 had been recovered recently by U.S. agents.

Should academics publish texts from cuneiform tablets that may have been looted? This thorny ethical question sparked the

fiercest debate at the meeting and revealed a bitter split within the community. Some philologists say that given the scale of the looting, they are eager to salvage what data they can by translating and publishing texts. “You have an obligation to your science, to your data,” says Jerrold Cooper, a philologist at Johns Hopkins University in Baltimore, Maryland, who says he would work with collectors who own tablets. “It makes no sense at all to condemn all publication” of potentially looted items.

But many archaeologists see the widespread looting in Iraq as an unalloyed nightmare and any involvement with potentially stolen tablets as aiding and abetting the destruction. At the meeting, a faction led by Michael Mueller-Karpe, a specialist in ancient metals at the Roman-German Central Museum of Mainz, Germany, proposed a res-

olution opposing scholarly involvement with tablets that may have been looted. “Scholars ... are urged to refrain from providing expertise to the antiquities market and to private collectors, unless the artifacts in question can be proved to be neither excavated illegally nor exported without permission,” states the resolution, which was signed by 130 academics at a meeting after the conference officially ended. A number of scholars, primarily philologists like Cooper, refused to sign.

The different opinions do not always track disciplinary lines. Robert Adams, a retired archaeologist and former head of the Smithsonian Institution, surprised many participants at the opening session by allowing that no discipline should be expected to ignore vast amounts of new data, however it might have been obtained. (After taking fire from colleagues, Adams later clarified that he did not mean to condone the publishing of looted material but wanted to emphasize the complexity of the problem.)

Meanwhile, several philologists draw a distinction between working on existing collections and trafficking with dealers seeking to boost the value of tablets. Cooper, for example, says he would “not be comfortable” examining tablets owned by dealers.

But a few at the meeting do read recently acquired tablets for dealers, for free or for pay—an act that archaeologists maintain can boost the tablets' value and reinforce the cycle of looting. Cooper says he hopes participants at the next conference will come up with a common ethical stance to guide scholarly actions.

—ANDREW LAWLER

Drugs, Quarantine Might Stop A Pandemic Before It Starts

Thirty-six years after the last influenza pandemic, researchers wonder whether they can make these global disasters a thing of the past

It might just work. With military-style planning, a big stash of pills, and a lot of luck, the world might be able to stop a nascent influenza pandemic dead in its tracks, two new modeling studies conclude.

The models, published online this week in *Nature* and *Science* (www.sciencemag.org/cgi/content/abstract/1115717), are the first attempts to estimate the power of the antiviral drug oseltamivir to quash a pandemic—an unprecedented and audacious idea. If large numbers of people in the region first hit by a pandemic virus take the drug prophylactically and comply with some quite draconian measures to limit their movements and contacts, millions of lives might be saved, the authors of both papers say—and medical history would be rewritten in the process.

But just how likely is that scenario to succeed? As experts point out, the models, both of which chose Thailand as the presumptive ground zero, are based on several untested assumptions: that the runaway virus isn't highly infectious, for instance, and that large quantities of drugs can be distributed rapidly to the right people, even in remote villages. "The models make sense, and we should seriously consider this approach," says Harvard epidemiologist Marc Lipsitch, "but the take-home message is there's no way we can count on this."

The researchers—one team led by Ira Longini of Emory University in Atlanta, Georgia, the other by Neil Ferguson of Imperial College London—hope their work will lead to con-

crete actions because until now, there's been little if any official commitment to such a plan. The World Health Organization (WHO), which the researchers say would have to lead the effort, is "interested," says the agency's pandemic chief, Margaret Chan. Rich countries are stockpiling oseltamivir to protect their own populations, but they have no plans yet for shipping it to the cradle of a pandemic. Nor are the Asian countries affected by H5N1—the avian influenza strain most feared as the potential source of the next pandemic—on board or necessarily up to the logistics, although they were slated to discuss the idea at a meeting in Bangkok earlier this week.

Given influenza's history, most experts peg the chance that the world will be hit by another pandemic at 100%. The question is when it will occur and how bad it will be; there's widespread agreement that the death toll could be in the tens of millions. Vaccines offer by far the best chance to avert that danger—at least in countries that can afford them—but these would take months to produce after a pandemic begins (*Science*, 15 October 2004, p. 394).

A bold new idea is to use oseltamivir to battle a potentially pandemic virus at the source, before it becomes a global threat, using an internationally run stockpile. The strategy might be

the only way to prevent disaster in the majority of countries unable to afford vaccines or drugs at all, notes Arnold Monto of the University of Michigan, Ann Arbor. Oseltamivir would make those who get the flu less infectious to others, but by far its most important task would be to prevent infection in those exposed to the virus.

Now, that idea has been put to the test. Longini and his colleagues simulated an imaginary population of 500,000 people who live, work, and move about in rural Southeast Asia. Meanwhile, Ferguson and his colleagues built a model based on the 85 million people living in Thailand and a 100-kilometer-wide border zone in neighboring countries. Both then introduced a pandemic virus and looked at how well different containment strategies performed.

The cornerstone in each model was giving a 10-day prophylactic course of oseltamivir to the contacts of every suspected flu patient—either by treating everyone in their household, school, or workplace, or by simply giving it to anyone living within a certain radius. In both models, the drug regimens were supplemented by measures such as closing schools, "home quarantine," or "area quarantine," in which travel into and out of the hot zone is restricted.

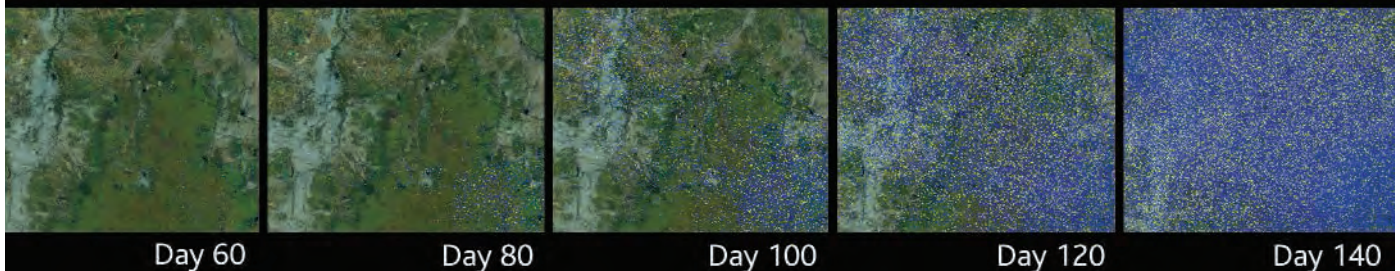
And in both models, the more such measures were deployed, the higher the chances were that the pandemic petered out, with thousands or even millions of people taking an oseltamivir course, but only a few hundred actual flu cases. But success depended critically on a few factors.

One is the infectiousness of the pandemic virus. Epidemiologists characterize infectious agents by a factor called R_0 , which denotes the number of secondary infections caused by a primary case. In both studies, viruses with an R_0 between 1.0 and 1.8 could usually be contained, depending on the exact set of measures; with an R_0 well above 2.0, the outbreak often spiraled out of control. Estimates for R_0 during past pandemics have varied; in a paper published in December, Lipsitch concluded that it was between 2 and 4 in the United States during the 1918–19 pandemic. But Ferguson estimates it was about 1.8.



Drug of the day. A global stockpile of up to 3 million treatment courses of oseltamivir might be needed.

Avian Flu Epidemic Without Intervention



Stop right there. A computer model shows how a new pandemic might spread in Thailand. With massive prophylaxis, the outbreak might end after just a few hundred cases.

CREDITS (TOP TO BOTTOM): ASSOCIATED PRESS; I. LONGINI ET AL.

A Drug Makes It Big—But Can It Deliver?

The worldwide fears triggered by the Asian outbreak of H5N1 have created one clear winner: oseltamivir, the drug that, from a quartet of candidates, is considered the best one to fight a pandemic. More than two dozen governments have placed orders for a stockpile with the producer, Roche in Switzerland; 2005 sales are expected to exceed \$700 million—up from just \$110 million 3 years ago—and seem poised to grow further, says Bret Holley, an analyst at CIBC World Markets in New York City.

The procurement orders may be lucrative, but it remains to be seen just how effective oseltamivir, known commercially as Tamiflu, will be during an influenza pandemic. Nor is there agreement about how big a national stockpile should be, or who should receive the drugs to maximize their impact. And in the worst-case scenario, resistance in the flu virus might render stockpiles worthless.

As a remedy against nonpandemic flu, oseltamivir has certainly failed to win many supporters since its launch in 1999. The drug, which blocks a viral enzyme called neuraminidase, can make a bout with influenza more bearable and shorten the duration of symptoms by a day or more; it has also been shown to prevent complications and hospitalizations—but not mortality. The problem is that it needs to be given within 48 hours of infection to be fully effective. And even for patients who meet that deadline, most doctors don't think the benefits warrant the \$65 cost of a prescription. (Japan, where sales have soared, is the exception.)

How well oseltamivir will perform against human infection with H5N1 is unclear. It has shown anti-H5N1 activity in test-tube and animal studies, but human cases have been so rare that experience is extremely limited.

Who should get treatment is also in question. Pandemics may sicken between 25% and 50% of the population in 3 months, but many people with milder cases can probably recover by themselves. Still, countries such as France, the United Kingdom, and Finland are amassing enough oseltamivir to treat 20% to 30% of their populations; the United States,

on the other hand, currently has only 2.3 million doses for almost 300 million people. The Bush Administration was expected to announce a new order shortly—although nowhere near the 67 million to 124 million treatments that the Infectious Diseases Society of America has urged.

Recently, another potential role of oseltamivir has garnered a great deal of attention: that of preventing illness rather than treating it. Studies have shown that oseltamivir can reduce the risk of infection in people exposed to the virus by around 80%. That benefit is key in global plans to stamp out a pandemic early on (see main text); once a virus is on its worldwide rampage, national governments could similarly attempt to slow its spread within their own borders.

Last year, a study by Ira Longini's team at Emory University in Atlanta, Georgia, showed that using oseltamivir preventively could contain an outbreak in the United States, and a paper published this month by Ran Balicer of Ben Gurion University of the Negev in Be'er Sheva, Israel, suggests that stockpiling drugs for this purpose should be cost-effective if pandemics occur more often than once every 80 years. That may seem like a fairly safe bet, but it would require reserves for much more than 25% of the

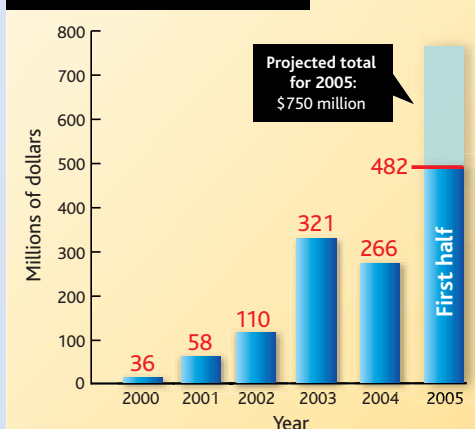
population—an amount few countries are considering at the moment.

For now, a more feasible and widely discussed approach may be to restrict prophylactic use to certain groups, such as health care workers, people performing "essential" jobs, or the elderly—although picking the beneficiaries might create wrenching ethical dilemmas.

Another worry is that once tens of millions of people start taking Tamiflu, the virus will become resistant. So far, resistance appears to be rare in other flu strains; during the 2003–04 flu season, when a whopping 6 million treatment courses were prescribed in Japan, only 4 of 1180 virus isolates tested there showed resistance, a group reported in April. And fortunately, mutations that confer resistance also appear to slow the virus's growth.

Tamiflu may soon face some competition, as other drugs are in the pipeline. And many researchers say they'd feel a lot better if the bullish market for flu drugs were split between a couple of rivals. —M.E.

Oseltamivir Global Sales



Another key condition in both models is that the operation starts within a couple of weeks of the first cases. Chances of containment drop dramatically if it takes more than 2 days to reach new patients' contacts. Both conditions may be challenges, to say the least, in rural areas with poor health care.

Some infectious-disease experts put little stock in models like these. "In 30 years in public health, I've never seen any statistical modeling that had any impact on public health. And this is no exception," says Michael Osterholm, director of the Center for Infectious Disease Research and Policy at the University of Minnesota, Twin Cities. A single SARS patient in a Hong Kong hotel triggered a worldwide outbreak in 2003, he notes; no model could have predicted that turn of events.

But to Anthony Fauci, director of the U.S. National Institute of Allergy and Infectious Diseases, the studies provide an "interesting

blueprint" of what might be possible. "Even if there's only a 20% or 30% chance of success, it's worthy of the effort," adds Frederick Hayden, an antiviral expert at the University of Virginia, Charlottesville, "given the enormous impact that a pandemic would have."

It wouldn't be all that expensive, Hayden notes. The amount of oseltamivir needed—some 3 million courses in Ferguson's most unfavorable scenario—isn't very much; the United Kingdom alone has ordered almost 15 million 5-day courses for its own citizens. WHO already has more than 100,000 treatment courses, donated by Roche, sitting in a stockpile. And Roche may soon make another, much larger donation to WHO, says David Reddy, the company's influenza pandemic task force leader.

But although WHO welcomes any oseltamivir it can get its hands on, more studies, as well as discussions with the affected

countries, are needed to find out whether the snuffing-out scenario is feasible, Chan says. The Thai government, for its part, is interested in exploring the option, says Supamit Chunsuttiwat, a senior expert for communicable diseases at the Ministry of Public Health. The two papers, he says, "give us some hope that we might be able to do this."

But Osterholm worries that the two papers might calm fears prematurely. Even if the scheme envisioned by Ferguson and Longini were successful once, he said, it would need to be repeated as long as H5N1 is rampant in the bird population. Longini agrees. But who knows, he says, researchers might get better at it after the first time. And in any case, only a small region would be affected in every budding pandemic. "It's not like we're exposing the entire world to a fire drill every time," Longini says.

—MARTIN ENSERINK

RANDOM SAMPLES

Edited by Constance Holden

Plague of Toads Down Under

Cane toads, brought to Queensland from South America in the 1930s to tackle sugarcane beetles, have become Australia's latest big alien menace. The toads have been gradually spreading across tropical northern Australia, and citizens of Darwin are bracing for an invasion when the next wet season starts in December.

The highly fecund cane toads (*Bufo marinus*) are a threat to native wildlife because of a poisonous sac at the base

of the head that can kill any predators, from crocodiles to cattle, that take a nip at them. They have been implicated in the recent drastic decline of quolls, cat-sized marsupials. Authorities worry that they will cut a swath through native frogs, snakes, and goanna lizards in the Northern Territory's famous Kakadu Park.

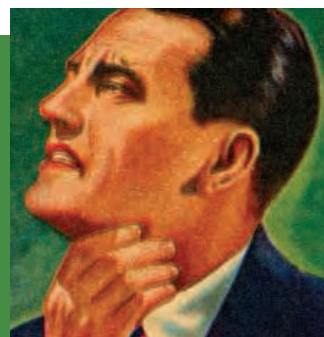
Trapping, monitoring, and public education activities are now in high gear. Zoologist Gordon Grigg of the University of Queensland in Brisbane and others are looking for "toad-specific" solutions, including scents and sounds that will draw



Adieu, Antibiotics

In the latest step to wean its citizens of their antibiotic habit, France has banned 12 popular sore throat remedies that contain "topical antibiotics." The drugs—lozenges, mouthwashes, and sprays—haven't been shown to do any good, the French Health Products Safety Agency has concluded. Most throat infections are impervious to antibiotics because they are caused by viruses, they note. And the drugs, all but one available over the counter, could foster the development of drug-resistant microbes.

Since 2001, the French government has campaigned to drive down the country's use of antibiotics, among the highest in Europe, which has resulted in high rates of treatment resistance in pneumonia and hospital-acquired infections. The government prohibited the use of several medically important antibiotics in sore throat remedies in 2003. The latest ban affects four others whose overuse posed a much smaller risk to public health, says infectious diseases specialist Stephan Harbarth of the University of Geneva in Switzerland. Still, he adds, the move signals to doctors and patients "that France is getting tougher."



Mal de gorge. It's time for Frenchmen to tough it out.

the animals to traps, pheromones that might be used to disrupt breeding cycles, and toxins special to the cane toad. In the long run, government scientists hope that genetic engineering will allow them to throw a wrench into tadpole development.

That solution is still far away, says Adelaide University zoologist Michael Tyler. Amphibians are rarely invasive species, he notes, so scientists still know little about

the toad and its potential Achilles' heel. That is why, he says, "everything we've tried so far has been unsuccessful."

Burning Boats for Animal Rights

Last month, animal-rights activists again attacked Oxford University over an animal research facility it is building, this time setting fire to a boathouse.

A year ago, activists used threats and property damage to force the main contractor to withdraw from the \$33 million project (*Science*, 23 July 2004, p. 463). Now the Animal Liberation Front (ALF) has claimed responsibility for the 4 July fire, which destroyed 24 shells and caused almost \$1 million in damage. No arrests have yet been made.

ALF has threatened the university and its suppliers with further attacks. "[N]othing ... is off limits until the project is scrapped," it wrote on its *Bite Back* Magazine Web site. A university spokesperson says it is still determined to resume work on the lab, designed mainly to house rodents.

Some animal-rights protesters say their tactics are a response to stiffer laws against economic sabotage. A new law allows up to 5 years' imprisonment of anyone who harasses businesses supplying research organizations (*Science*, 4 February, p. 659).

Bombs Away ... and Back

After a 4-year absence from the Bradbury Science Museum at Los Alamos National Laboratory, the Little Boy nuclear bomb has returned—sort of.

Shortly after the 11 September 2001 terrorist attacks, the museum and its counterpart at Sandia National Laboratories removed their Little Boy bombs, 4500-kilogram cousins of the device that killed more than 70,000 people at Hiroshima, Japan, 60 years ago tomorrow. Although the bombs lacked nuclear or explosive material, says Bradbury director John Wheaton, the government worried that their "anatomically correct" innards might interest terrorists. "These old, crude, brute-force World War II devices [had] suddenly become an issue," he says. (The museum's Fat Man, a copy of the one used at Nagasaki, is only a bomb case.)

Wheaton says Sandia carried out "surgery" on the private parts of its Little Boy. (Sandia would not confirm this.) But he says Bradbury's bomb was transferred to a well-guarded facility while modelmakers in Texas built the replica of its casing now on display.



Little Boy now just a toy.

Edited by Yudhijit Bhattacharjee

JOBS

Stanford bound. Four decades of research and administration at Los Alamos National Laboratory (LANL) has taught Siegfried Hecker a thing or two about nuclear weapons and proliferation. Now the 61-year-old metallurgist will get to share that knowledge with students at Stanford University in California.

Hecker, who served as LANL's director until 1997 and stayed on as a scientist before retiring from the lab last month, is going to Stanford's Center for International Security and Cooperation this fall as a visiting professor. There, he'll continue his work on curbing proliferation and nuclear terrorism and teach courses on science and nuclear security. Hecker says the looming new LANL contract clinched his decision to leave

the New Mexico lab.

"I got waylaid," he says, about an academic post he turned down in 1968 to join Los Alamos, where he became an authority on plutonium. "And now, 37 years later, I'm finally doing it." Center Director Scott Sagan says it's going to be "very exciting" to have him around.

Hard labor. Scientists made "a huge mistake" this spring by sitting out the Kansas school board hearings on intelligent design, says science historian Niall Shanks, who has battled creationism in debates and writings. But he says his decision to join the philosophy department at Wichita State University (WSU) in Kansas, ground zero for the creationism movement, has little to do with that fight. It's "not really what I'm interested in," says British-born Shanks, calling the defense of evolution "dirty work [that] someone's got to do."

It's an endowed position at Wichita that lured him from his

current post at East Tennessee State University, says Shanks, who may extend his work from the study of biological self-organization and complexity into the philosophy of medicine. WSU philosophy department chair David Soles says Shanks's teaching ability and academic achievement won him the job, not his activism, although he says he's "gearing up for some

flak" from the appointment. University of Kansas science education professor Steve Case takes umbrage at the criticism of the community's strategy by his new ally. "But otherwise I'm thrilled," he said.

Got any tips for this page? E-mail people@aaas.org

AWARDS

Sounds good. Masakazu Konishi and his former postdoctoral researcher, Eric Knudsen, will share the \$200,000 Neuroscience Prize from the Peter Gruber Foundation. Konishi, a professor at the California Institute of Technology in Pasadena, and Knudsen, a professor at Stanford University in California, receive the honor for their research on the neural circuits and mechanisms that underlie sound localization in barn owls. The prize will be awarded at the Society for Neuroscience's meeting in November.

DEATHS

A smoking gun. British epidemiologist Richard Doll, who helped save millions of lives by showing a causal link between smoking and lung cancer, died last week in Oxford, U.K., at the age of 92.

Working with Austin Bradford Hill, Doll began asking British doctors in 1951 what they smoked and documenting what they died from. The work confirmed that smokers were much more likely than non-smokers to die of lung cancer.

In 1969, Doll became a professor of medicine at Oxford University, where he teamed up with Richard Peto and showed that smoking could also cause other types of cancer, as well as heart disease and respiratory disease. Doll's research "has done as much to save lives as the discovery of penicillin or the development of the polio vaccine," says neuroscientist Colin Blakemore, chief executive of the Medical Research Council in London. "The profound implications for health policy resonate to this day."

IN THE NEWS

Magic lessons. Call it "Harry Potter Meets the Teacher of Science." Christine Redman of the University of Melbourne in Australia has been using the wildly popular series to harness students' interest in science and mathematics. Her Potter-centric curriculum employs flying broomsticks to reveal fundamentals of mechanical engineering and Dementor-curing chocolate to illustrate brain biochemistry and social behavior. Redman has trained more than 600 chemistry school teachers in the past 4 years, and the state of Victoria offers her material to all its schools. Redman says she plans to create more lessons based on the latest volume, a 672-page spellbinder, once she finishes reading it.



Scrapie in Ancient China?

SCRAPIE, THE TRANSMISSIBLE SPONGIFORM encephalopathy found in sheep, is generally traced back to the early 18th century, based on records in the UK, France, and Germany (1). However, I believe philological evidence suggests that the disease may have existed up to 2000 years earlier.

Chinese characters are often composed of two parts. One component may hint at meaning, while the other gives the sound, or both may contribute to the meaning. Since these characters were developed from 3000 to 2000 years ago, “original intent” can only be inferred. One character for “pruritis” or “itchy” is composed of three parts, an outer wrapper meaning “disease” and inner parts meaning “hand” and “insect”:

瘙 = 疒 + 又 + 虫
itchy disease hand insect

Another character for “itchy” combines “disease” and “sheep”:

痒 = 疒 + 羊
itchy, disease sheep
pruritis

The English name “scrapie” arose because afflicted sheep suffer pruritis and wind up scraping off much of their coats. I propose that this dramatic symptom was known in ancient China and hence the combination of “disease” and “sheep” to make “pruritis.” The importance of sheep at that time in China is suggested by the character for nutrition, 養, which combines 食, to eat, with 羊, sheep.

Chinese characters evolved over time, but this one has likely been unchanged for at least the past 1000 to 1500 years. A.D. 500 to 1000 is the era during which Japan adapted Chinese characters to the grammatically very different Japanese language. The character for “itchy” (kayui) in Japanese is identical to that used in traditional Chinese, indicating an ancient origin for this form of the character. Both 痒 and 瘙 are pronounced “yang” in Mandarin Chinese and “yō” in Japanese, so it is possible that 痒 is giving the sound. However, many disease-related characters seem to get both meaning and sound from the enclosed part:

知 (to know, “chih”)
痴 (stupid, “chih”)
包 (a packet, “pao”)

疱 (blister, “pao”)
古 (old, “ku”)
疔 (chronic complaint, “ku”)
豆 (bean, “tou”)
痘 (smallpox, “tou”).

The prion model for scrapie suggests that it should occur spontaneously at some low frequency in all sufficiently large herds. The remarkable stability of the scrapie agent suggests that infectivity may even be detectable in remains of ancient sheep.

REED B. WICKNER

Laboratory of Biochemistry & Genetics, National Institute of Diabetes and Digestive and Kidney Diseases, NIH, 8 Center Drive, MSC0830, Bethesda, MD 20892–0830, USA. E-mail: wickner@helix.nih.gov

Reference

1. H. B. Parry, *Scrapie Disease in Sheep—Historical, Clinical, Epidemiological, Pathological and Practical Aspects of the Natural Disease* (Academic Press, London, 1983).

Co-Funding in Canada: Another View

THE LETTER “PROBLEMS WITH CO-FUNDING in Canada” by M. Tyers *et al.* (24 June, p. 1867) gives a misleading and inaccurate picture of recent increases in federal research funding. The Letter leaves the distinct impression that

Canadian scientific excellence is being compromised by an overemphasis on short-term industrial and commercial interests. Statements such as “[m]uch of this renewed commitment to research... is in the form of “co-funded” programs...” and “co-funding is often biased against fundamental research...” not only belie the facts but do a

tremendous disservice to a range of innovative Canadian programs that have changed the face of the research environment in Canadian universities over the last 7 years and have reaped a “brain gain” of exceptional young and established stars from other countries. Federal allocations to university R&D have increased by \$13 billion Canadian since 1997–98, giving Canada the highest expenditures on higher education R&D per capita among G8 countries. For 2003–04 and 2004–05, only 22% of total expenditures from all federal granting bodies required co-funding. The authors fail to mention the doubling of the federal granting council budgets; the creation of the 2000 Canada Research Chairs; Millennium and Canada Graduate Scholarships; permanent funding for the

Networks of Centres of Excellence; and the funding of indirect costs of research.

Although co-funding programs represent a relatively small portion of the total federal investment, they are an important element in a range of funding vehicles designed to foster a world-class research enterprise in Canada. The largest programs with co-funding requirements are the Canada Foundation for Innovation (CFI) and Genome Canada (GC). To date, CFI, through its 60% matching requirement, has generated \$4.3 billion in co-funding for university infrastructure from \$2.7 billion in federal investments. For the much smaller GC program (which enables Canada to compete internationally in large-scale genomics research projects), co-funding has leveraged an additional \$470 million for genomics researchers from a \$375 million federal investment. Genomics programs in the Netherlands and Spain have been modeled after GC. Similar programs in Sweden, UK, and the United States use a co-funding model.

For both the CFI and GC programs, provincial governments and not industry have provided most of the co-funding, a fact not mentioned by the Letter writers. Co-funding is an important component of funding in many countries, particularly for very large projects, to harness resources that would not

“ [C]o-funding programs...are an important element in a range of funding vehicles designed to foster a world-class research enterprise in Canada.”

—CARTY

otherwise be accessible to researchers.

Another major point made in the Letter is that “[by] eschewing scientific excellence as the primary consideration, co-funded programs imperil scientific credibility...,” implying that co-funded proposals fail to meet the highest standards of excellence set by international peer review. This is demonstrably not the case: For both CFI and GC competitions, review panels have been made up of individuals with impeccable scientific credentials. Publicly funded research in Canada is subject to rigorous peer review, and scientific excellence is not compromised. However, scientific merit is not necessarily the sole determinant of success. For projects involving multimillion dollar investments of public money, funders have a responsibility to ensure that grantees are

accountable for the use of funds through sound financial and management practices.

In conclusion, I believe the facts show that Canada and Canadian researchers have benefited enormously not only from the massive federal investments, but also from the additional \$5 billion that has been leveraged via co-funding mostly from provincial governments over the past 7 years and from the broader engagement of the R&D community. However, a portion of publicly funded research is subject to a double test of peer-reviewed excellence and the ability to attract co-funding partners from the public and private sectors. The result is more support for science and a stronger underpinning for the science and technology enterprise in Canada.

ARTHUR J. CARTY

National Science Advisor to the Prime Minister, Government of Canada, Blackburn Building, Room 728, 85 Sparks Street, Ottawa, ON K1A 0A3, Canada. E-mail: acarty@pco-bcp.gc.ca

Update on the Closure of a Chilean Institute

IN REFERENCE TO THE LETTER "THE END OF A Chilean institute" (L. Barbeito *et al.*, 6 May, p. 792), we confirm that the Board of

Directors of the Chilean Millennium Science Initiative Program (MSI) unanimously resolved to discontinue funding for a second 5-year period for one of its three science institutes, the Institute for Studies in Cell Biology and Biotechnology (CBB). This decision was based on the recommendation of the MSI Programme Committee, composed of eight prestigious foreign scientists (five of whom are members of the U.S. National Academy of Sciences) who are entrusted with the selection and evaluation of the MSI Centers of Excellence (www.mideplan.cl/milenio). The basis for the decision was a deep split between the Director and the six lead scientists of the institute, caused by differences in research objectives, approaches, and internal procedures. Although acknowledging CBB's scientific achievements, the MSI seeks and expects responsible, professional, and efficient management of its resources and personnel.

The MSI Board of Directors decided to continue supporting students and postdoctoral fellows previously attached to the CBB. At the same time, a call for proposals to create a new institute to replace the gap left by CBB was announced on 17 April 2005 and is open to interested groups of scientists in all scientific fields, including previous CBB scientists.

In the last 5 years, the government of Chile and the National Commission for Scientific and Technological Research have increased investment in science and technology by ~60% and the number of doctoral grants by fivefold. With its three Institutes and 15 Science Nuclei, the Chilean MSI will continue its mission of significant contributions to scientific and technological development, education of young people, and international cooperation.

YASNA PROVOSTE¹ AND ERIC GOLES²

¹Chilean Minister of Planning, Chairwoman Board of Directors Millennium Science Initiative, Ministerio de Planificación, Ahumada 48 piso 10, Santiago, Chile. E-mail: icm@mideplan.cl. ²President, National Commission for Scientific and Technological Research (CONICYT), Canadá 308, Santiago, Chile. E-mail: secrepriv@conicyt

Canadian Database in Singapore

I AM WRITING IN RESPONSE TO THE ARTICLE "Canadian global database may move to Singapore after loss of funding" (News of the Week, W. Kondro, 10 June, p. 1529). I would like to clarify some statements made in the article that I feel misstate the current situation.

First, Blueprint and its databases are not “moving” to Singapore, because we are already in Singapore. The Blueprint Initiative Asia Pte. Ltd. (Blueprint Asia), the mirror node of Toronto’s Blueprint North America, is a nonprofit company registered in Singapore in 2004 that has been a vital cog in the organization’s efforts to develop the Biomolecular Interaction Network Database (BIND). With a staff of 15 curators and two administrators, Blueprint Asia focuses on human diseases of interest to researchers in Asia and supports biotechnology, bioinformatics, and human resources development in Singapore.

Second, the photo caption seems to imply that the Blueprint layoffs only affected staff involved in curation activities. Among the staff released on 2 May were several highly prized software developers responsible for systems such as BIND, the SeqHound data warehouse, and data visualization technologies.

Blueprint is quite excited about the prospect of expanding our operations in Singapore, which has long been part of our plan. We have found the Singaporean community to be quite welcoming of our efforts, recognizing the long-term global implications of our undertakings, and they have been extremely supportive, scientifically, socially, and financially. Our only regret is that this

good news was prompted by the possible closure of Blueprint’s Toronto operations.

CHRISTOPHER HOGUE

Principal Investigator, The Blueprint Initiative, 522 University Avenue, 9th Floor, Suite 900, Toronto, ON M5G 1W7, Canada.

Random Copying and Cultural Evolution

ALTHOUGH NETWORK THEORY HAS MUCH TO offer, the mathematical study of collective human behavior is older and richer [e.g., (1–4)] than Albert-László Barabási suggests in his Perspective “Network theory—the emergence of the creative enterprise” (29 Apr., p. 639). Barabási rightly makes the point that, rather than studying static networks, the time has come to study how social networks evolve in time, as some have already begun to do (5–7). In laying out this objective, however, Barabási curiously does not include evolutionary theory, which is specifically the study of how the frequencies of variants change over time. In fact, Cavalli-Sforza and Feldman (1) demonstrated decades ago how the mathematical theory of population genetics has all the tools to study change in human behavior over time. A powerful tool is the neutral

theory of random genetic drift (8), by which a population of individuals copy variants from each other, except for a small fraction in each time step who invent a new variant (9–11). Many of the phenomena addressed by network theory are also elegantly explained by random copying (7, 8).

As a null hypothesis, the random-copying model is simpler than network theory. Are Web links, Hollywood actors, and scientific collaborators really nodes in a network, or are they just ideas that are copied among individuals? In the Barabási-Albert model (12), “preferential attachment” is an imposed rule, whereas in the random-copying model, the “rich get richer” effect emerges naturally because the more popular a variant is, the more likely it will be copied again, becoming even more popular. Time-dependent difficulties for the basic preferential attachment network model (13), such as quick success by new nodes and rapidly changing networks, are not a problem in the random-copying model, by which any new variant stands a small chance of becoming highly popular (like a network “hub”), and the network of who copied whom completely changes at every time step. Given these advantages, combined with decades of research establishing neutral theory (14), network researchers

should broaden their connections outside their own emerging and exciting field.

R. ALEXANDER BENTLEY^{1*} AND
STEPHEN J. SHENNAN²

¹Department of Anthropology, University of Durham, Durham DH1 3HN, UK. ²AHRB Centre for the Evolutionary Analysis of Cultural Behavior, University College London, 31-34 Gordon Square, London WC1H-0PY, UK.

*To whom correspondence should be addressed.
E-mail: r.a.bentley@durham.ac.uk

References

1. L. L. Cavalli-Sforza, M. W. Feldman, *Cultural Transmission and Evolution: A Quantitative Approach* (Princeton Univ. Press, Princeton, NJ, 1981).
2. R. Boyd, P. J. Richerson, *Culture and the Evolutionary Process* (Univ. of Chicago Press, Chicago, IL, 1985).
3. S. J. Shennan, *Genes, Memes and Human History* (Thames & Hudson, London, 2002).

4. C. Renfrew, K. L. Cooke, *Transformations: Mathematical Approaches to Culture Change* (Academic Press, New York, 1979).
5. R. A. Bentley, H. D. G. Maschner, *Fractals* **8**, 227 (2000).
6. R. A. Bentley, S. J. Shennan, *Am. Antiq.* **68**, 459 (2003).
7. P. S. Dodds, D. J. Watts, *J. Theor. Biol.* **232**, 587 (2005).
8. M. Kimura, J. F. Crow, *Genetics* **49**, 725 (1964).
9. F. D. Neiman, *Am. Antiq.* **60**, 7 (1995).
10. M. W. Hahn, R. A. Bentley, *Proc. R. Soc. London Lett. B* **270**, S120 (2003).
11. R. A. Bentley, M. W. Hahn, S. J. Shennan, *Proc. R. Soc. London B* **271**, 1443 (2004).
12. A.-L. Barabási, R. Albert, *Science* **286**, 509 (1999).
13. R. Albert, A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
14. J. H. Gillespie, *Population Genetics, A Concise Guide* (John Hopkins Univ. Press, London, 1998).

An Unexpected Cover Image

CONGRATULATIONS TO SCIENCE ON WHAT must be your new campaign to generate interest among nonscientists! For the first time ever, my husband picked up an issue of *Science* (Special Issue on Women's Health, 10 June). Unfortunately, he did not open the magazine, but he did turn it over to see if the nude model on the front cover (posterior view) appeared in an anterior view on the back cover.

My observational study ($n = 1$) clearly proves that although sex may sell, it doesn't compute. Please return to electron photomicrographs and other such scientific

esoterica on the cover, for those of use who sit down with *Science* for a few moments of pleasurable reading at the end of a long day.

LINDA FELVER

Oregon Health & Science University, Portland, OR 97239–2941, USA.

CORRECTIONS AND CLARIFICATIONS

Reports: "Amplification of acetylcholine-binding catenanes from dynamic combinatorial libraries" by R. T. S. Lam *et al.* (29 Apr., p. 667). Reference 1 is incorrect. It should be: P. A. Brady, R. P. Bonar-Law, S. J. Rowan, C. J. Suckling, J. K. M. Sanders, *Chem. Commun.* **1996**, 319 (1996).

Research Articles: "A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*)" by S. A. Goff *et al.* (5 Apr. 2002, p. 92). Syngenta's rice genome sequence information was donated to the International Rice Genome Sequencing Project (IRGSP) and the Beijing Genomics & Bioinformatics Institute (BGI) in 2002 under an agreement to integrate the Syngenta data with the sequence generated by these projects and deposit the assemblies into appropriate public databases such as GenBank and EMBL. The individual sequence reads from Syngenta were deposited into the GenBank Trace Database by the BGI in February 2004. The Syngenta draft sequences can be obtained by request at the Web site www.tmri.org/en/partnership/access.aspx, although the authors recommend the use of the more complete sequence now in GenBank.

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 6 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

Sumptuous Survey of Hexapod History

Edmund Jarzembowski

If the number of described species is a measure of success, then insects (with nearly 1 million) are the most successful group of all time. And if the number of families known from the fossil record is a proxy for past biodiversity, then insects (with more than 1200) are also the most diverse paleo group. Insects have only existed for

11% of the duration of life on Earth. Yet in that time, they have pervaded all terrestrial ecosystems and evolved social organization several times. They conquered the air long before any flying vertebrate and

have outlasted trilobites and dinosaurs. They pollinate our crops and arguably gave us our greatest laboratory animal, *Drosophila*. Love them or hate them, we have evolved alongside them.

There are a number of good entomology books on the market. Few, however, have integrated the living and fossil record as seamlessly as David Grimaldi and Michael Engel's *Evolution of the Insects*. None, moreover, has combined this integration with so much student-friendly text and such a wealth of illustrations (more than 900). The book shows that lavish photography and lucidity need not be the prerogative of popular entomology and that segregation of entomology and paleoentomology is tantamount to intellectual apartheid. While looking good, *Evolution* is no coffee-table adornment. Weighing in at 2.92 kg, it is the western challenger to *History of Insects (I)*, a multiauthored, English-language account by Russian paleoentomologists. That work, which highlighted the fossil insect riches of Asia, is unashamedly phyletic; its approach to evolutionary relationships combines cladogenesis and evolutionary divergence. *Evolution* offers a cladistic treatment rooted in the Hennigian tradition—as one would expect from the previous contributions of Grimaldi (a curator of invertebrate zoology at the American Museum of Natural History) and Engel (a paleoentomologist at the

University of Kansas). This methodological difference influences the interpretation of the evolutionary history of insects (e.g., the longevity of groups like caddisflies and cockroaches).

One should never judge a book by its cover, but it's a fine fossil that fronts the volume—and the exceptionally preserved longhorned grasshopper in goethite is one of many superb compression fossils illustrated in full color. Cretaceous and Tertiary amber inclusions are also expertly photographed, but some depictions of internationally important fossil-insect localities are less successful. Many specimens look better than in their original publications. Images of fossils are liberally interspersed with those of living insects and even other arthropods, and numerous black-and-white scanning electron micrographs show fascinating detail, as only they can. Enhancement is mercifully kept to a minimum (I spotted only one tinted figure). Unfortunately, the captions do not always give



Fine fossil. This orthopteran (of the extinct family Elcanidae) is one of some 300 species (representing approximately 100 families) known from the Early Cretaceous Santana formation of Brazil.

information like species or scale. The former omission is at least salutary as many species have yet to be described scientifically. Phylogenies (which, the helpful glossary notes, are used in the book in the restricted sense of cladograms “superimposed on a geological time scale”) are accompanied by tables that identify key characters or organisms. The volume emphasizes morphology (especially the form of adults), but some behaviors are included, notably those suggested by trace fossils and color patterns (extant grasshoppers and butterflies). The authors, however, reject evidence for pre-Cretaceous social insect activity, and, curiously, they do not illustrate structural colors from fossil insects (such as German Eocene beetles).

Some slips are inevitable in a work of this scope. Among the errors I noted: The Bembridge Marls' insects are from earliest

Oligocene (not Late Eocene) limestones, which yield the butterfly *Nymphalites zeuneri* but not leaf mines; *Sinonemestrius* crops up in two different places in the phylogeny; the Purbeck-Wealden deposits are latest Tithonian–earliest Aptian; and the trigonotarbid in figure 3.7 is really a phalangiotarbid. The index omits some taxa, but typos and taxonomic changes are fortunately few—although the suggestions of new “popular” names for giant dragonflies (griffenflies) and extinct cockroaches (roachoids) lack luster.

The meaty text is divided into 15 substantial chapters. The first reviews biological basics (the nature and number of species, and reconstructing evolutionary history), and the second focuses on geology (insect fossilization, dating, and major deposits of fossil insects). The authors acknowledge the systematic value of genetics but note that “it is highly unlikely that any DNA is preserved in ancient fossils.” Nevertheless, I suspect that the search will continue, and the geological time scale used by the authors has already been supplanted (2). Grimaldi and Engel then turn to the insects' place in the phylum Arthropoda: it's evidently closer to millipedes than to crustaceans, although insects made landfall separately. Regrettably, they provide only a reconstruction of the “most famous Devonian hexapod,” *Rhyniella praecursor*.

The authors begin their treatment of the true insects (ectognathans, that is, those with external jaws) with an overview of the group's natural history, morphology, and phylogeny. They devote a short chapter to the earliest insects such as the bristletails, silverfish, and the early Devonian *Rhyniognatha* (which, although more derived, is considered the oldest known true insect). The primitively-winged insects (Pterygota) are covered in nine chapters, which proceed from the mayflies, extinct beaked insects, and odonatopterans (dragonflies and their extinct relatives) through the holometabolous orders (insects with a pupa), which attained awesome hyperdiversity. We still do not understand exactly how wings arose, and the mystery remains why no wings are found until the mid-Carboniferous given that basal (primitive) Pterygota include aquatic orders that inhabit depositional environments.

In the chapter outlining the course of insect evolution through the Cretaceous (K) and Tertiary (T) periods, the authors conclude “that the mechanism is obscure as to how insects diversified with angiosperms.” Mass extinctions and the

The reviewer is at the Maidstone Museum and Bentlif Art Gallery, St. Faith's Street, Maidstone, Kent ME14 1LH, UK. E-mail: edjarzembowski@maidstone.gov.uk

K/T boundary are only briefly considered, because Grimaldi and Engel distrust taxic evidence (counting records) and trust ghost ranges (cladistic extrapolation). They do note that many living insect species date back to the Tertiary and that insects are climatically sensitive and mobile. Thus biogeographers must beware—Gondwanan insects can lack austral ancestry.

In the short epilogue, the authors briefly review possible explanations for the bewildering number of insect species: the clade's age, morphology, capacity for high speciation rates, and low rates of natural extinction. Then they end with a couple of pages of comments on current threats and the future. Insects face mass destruction in the "6th mass extinction" now under way, and they need proselytizers now. Whatever is in store, *Evolution of the Insects* superbly documents the rich and colorful history of hexapods.

References

1. A. P. Rasnitsyn, D. L. J. Quicke, Eds., *History of Insects* (Kluwer Academic, Dordrecht, Netherlands, 2002).
2. F. M. Gradstein, J. G. Ogg, A. G. Smith, Eds., *A Geologic Time Scale 2004* (Cambridge Univ. Press, Cambridge, 2004).

10.1126/science.1115258

NATURAL HISTORY

A Swarm of Bee Books

Jay Hosler

In the introduction to *The Hive*, Bee Wilson confesses that she has "always felt fond of bees—or at least fond of the idea of them." As it turns out, she isn't alone. In the last year or so, there have been several popular books released that examine the place of bees in the ecosystem of human thought and culture. In Wilson's book, we see bees through the eyes of a food writer and historian of ideas who begins with the premise that "studying bees is a way of studying ourselves." Throughout human history, our understanding (or in many cases, misunderstanding) of honeybees has been a Rorschach test for our own cultural and institutional beliefs and ideas. The three books reviewed here approach the topics of honeybees and beekeeping from three distinctive perspectives—with engaging results.

The reviewer is in the Biology Department, Juniata College, 1700 Moore Street, Huntingdon, PA 16652, USA. E-mail: hosler@juniata.edu



Standing in front of a hive can stimulate a number of emotions. To the uninitiated, the most powerful is often the desire to beat a hasty retreat. Yet at one point in time even the most experienced beekeepers have approached a hive for the first time, opened it to the roaring buzz of thousands of venomous bees, and resisted that urge to run. Why? For Holley Bishop the answer appears to have

been love. Bishop's *Robbing the Bees* is the enthusiastic memoir of a writer who has recently come to beekeeping. Her romantic prose is veined with the poetic descriptions of someone who has just discovered an alien world and tempered by her practical experiences as a beekeeper. There are times when her descriptions of the natural history may be a bit too poetic, but such license is inevitable in a popular science book and forgivable given the rigor that permeates most of her account.

Each chapter (and the book as a whole) is framed by the story of a Florida beekeeper named Donald Smiley. We follow the affable Smiley on a journey through an entire honey season: chasing tupelo blooms, getting stung, extracting honey, going to market, and preparing for the next go-around. Along the way, Bishop stops to explore the history and economics of honey, wax, and beekeeping. Smiley's years of beekeeping experience provide an excellent counterpoint to Bishop's own relative inexperience, making the book as much about bees and honey as it is about Bishop's discovery of bees and honey.

Whereas beekeeping and honey are the main thrust of Bishop's book, *Letters from the Hive* by University of Arizona entomologist Stephen Buchmann (with writer Banning Repplier) and Wilson's *The Hive* spend far more time exploring the cultural and historical aspects of the bees. Buchmann's *Letters* is the passionate plea of a pollination biologist, bee enthusiast, and concerned conservationist who sees bees and ancient beekeeping traditions slowly disappearing. His writing stimulates our innate biophilia (E. O. Wilson's term for an unconscious love of living things) with tales such as of the Australian Aborigines' quests for the hives of stingless bees (called sugarbags) and arcane Mayan beekeeping rituals. Then, once we're drawn into the lives of these interesting people, Buchmann makes the experience bittersweet with descriptions of habitat destruction and lost cultural traditions. There are a few places in *Letters* where the writing succumbs to strained bee puns or cloying sweetness (as in the closing letter to the bees), but for the most part it is a fun book that briskly covers a lot of ground.

Buchmann is at his best when he is telling vivid stories like his experiences on a nocturnal honey hunt in the Kedah province of Malaysia. His guide was Pak Teh, the leader of a clan who harvests honey from a protected

Robbing the Bees
A Biography of Honey:
The Sweet Liquid Gold
That Seduced the World
by Holley Bishop

Free Press, New York, 2005.
334 pp. \$24, C\$35. ISBN
0-7432-5021-4.

Letters from the Hive
An Intimate History
of Bees, Honey, and
Humankind

by Stephen Buchmann
with Banning Repplier
Bantam, New York, 2005.
288 pp. \$24. ISBN 0-553-
80375-1.

The Hive
The Story of the
Honeybee and Us
by Bee Wilson

John Murray, London,
2004. 320 pp. £14.99. ISBN
0-7195-6409-3.

forest in the Pedu Lake region. (Teh has received written permission for his hunt every year since he petitioned the sultan in the 1960s.) The story is laced with elegant, ancient rituals and the fable of Hitam Manis, a beautiful Hindu handmaiden whose requited love for a sultan's son ultimately leads to her death and transformation into "the giant honey bees of the Asian rainforests." Through Buchmann's eyes, we watch Pak Teh and his clan climb a herringbone ladder up a giant tualang tree. There, on branches a hundred feet or more above the ground, the combs of the giant honeybees hang from the underside of branches like golden half moons. Using the falling sparks of a liana torch, Pak Teh tricks the giant bees off the comb long enough to cut it free from the tree with a traditional wood tool. All the while, Teh's brother-in-law stands at the base of the tualang tree singing the song of Hitam Manis.

Unlike Bishop and Buchmann, Wilson doesn't keep bees herself, but she is interested in honey, the idea of bees, and their relationship with humans. Of the three books, this is the most distinctive and compelling contribution. There have been numerous excellent books on bees and apiculture, but none to my knowledge examine the ideas of honeybees and beekeeping like *The Hive*. In an age of cheap sucrose, we have forgotten how much bees and honey have flavored our language and thinking. Wilson harvests thoughts from myriad sources and lays bare for the reader a secret history of the honeybee. Her wide-ranging book draws on thinkers from Democritus to Dickens and explores honeybees in terms of important topics like work and sex.

Wilson's chapter on politics is particularly enlightening. Our tendency to anthropomorphize the social structure of honeybees makes

these insects an alluring intellectual trap. Depending on who is peering into the hive, honeybees have been used as exemplars of monarchy, republics, autocracy, and dystopia. This tendency to read what we want into the hive is still with us. To favorably contrast bees with what she describes as "thuggish" wasps, Bishop calls bees pacifists who only attack when provoked—of course, the same can be said about wasps. As admirable and amazing as bees may be, it is always dangerous to instill them with too much humanity just because they make something as sweet as honey.

Despite the authors' different approaches, there is a fair bit of overlap among the three books. They all discuss the use of honey in cooking and include sections with recipes (I confess to skipping these). Each talks about mead and offers a different assessment about its worth as a beverage. They present historical and mythical anecdotes about bees and honey, and they cover beekeeping to a greater or lesser degree. All three authors discuss honey's role as an ancient folk remedy and the growing body of modern medical research that is confirming some of these traditional uses. Honey's hydrogen peroxide and extraordinarily high sugar concentration do in fact make it a good antibacterial agent. Both Bishop and Buchmann relate recent studies of burn victims treated with honey dressings who improved faster than patients receiving traditional chemical treatments. Thus we can add medicine to the long list of things we rob from bees.

At the end of *The Hive*, Wilson concludes:

However much human beings have projected themselves on to the hive, identifying themselves with drones, workers and the queen, and idealizing the morals of the waxen community, there will always remain mysteries about the life of the bees which men can never discover. And it is for this very reason that humans will continue to search for truths about themselves in the gold of the honeycomb.

If she's right, we can certainly look forward to more bee books in the future. I hope those will be as enjoyable and informative as these three.

10.1126/science.1115941



BROWSEINGS

A Dazzle of Dragonflies. Forrest L. Mitchell and James L. Lasswell. Texas A&M University Press, College Station, TX, 2005. 224 pp. \$39.95. ISBN 1-58544-459-6.

The authors offer an enticing introduction to the natural history and appreciation of dragonflies. This Mayan setwing (*Dythemis maya*) is one of the aerial predators whose color and variety are showcased in spectacular digital scans and field photographs.

CREDIT: GREG W. LASLEY FROM A DAZZLE OF DRAGONFLIES

Of Stones and Health: Medical Geology in Sri Lanka

Chandrasekara Dissanayake

To most people, geology is a subject concerning rocks and minerals and their distributions on Earth. As an earth science, of course, geology also deals with physical processes of the Earth, such as mountain building, coastal and river dynamics, and desert formation. Few people know that geoscientists have expanded their discipline by linking human and animal health with geology. The fundamental basis for this expansion, which has spun into a specialty known as medical geology, is the unique interdependence of the different living and nonliving components that make up the Earth. The basic building blocks of the Earth—the rocks and minerals—must, therefore, have a bearing on the health of the human and animal populations that live on these earthy materials.

Once a little-known specialty among a few geologists around the world, medical geology is now recognized as a field unto itself. Established by the International Union of Geological Sciences in early 1990, the Commission on Geological Sciences for Environmental Planning (COGEOENVIRONMENT) has defined medical geology as “the science dealing with the influence of ordinary environmental factors on the geographical distribution of health problems in man and animals.” It

now has developed into a truly fascinating new science, with potentially enormous consequences for the well-being of people around the world.

We Are What We Eat and Drink

On our planet, the chemical elements flow through the different planetary compartments, including the atmosphere, hydrosphere, lithosphere, and biosphere. Humans and animals are part of these cycles. The chemical elements pass into and out of them, too, in a complex biogeochemical cycle. Obviously, then, the chemistry of any local geological environment must have a direct influence on the chemical make-up of those living there. This is most readily seen in places where humans live in particularly intimate contact with the local physical environment, as is

the case with rural people living in tropical countries. Those living on lands with heavily impoverished soils, such as in Maputaland, South Africa, have such a low intake of essential elements that a very large percentage of the population suffers from a variety of diseases caused by severe mineral imbalances. Likewise, in other areas, there is an excess intake of elements due mainly to the abundance of certain elements in the environment. This leads to high incidences of mineral toxic-

ity, such as the widespread and tragic arsenic poisoning in India and Bangladesh.

Even though many other factors—among them life-style, sex, age, migrations, and food habits—affect health, imbalances in the supply of inorganic elements exert marked influences on both human and animal health. Anomalies in the local abundances of trace elements, for example, have a large impact on food chains. As it was more than 500 years ago, it remains relevant to bear in mind the basic law of toxicity as defined by Paracelsus (1493–1541), the father of pharmacology: “All substances are poisons; there is none which is not a poison. The right dosage differentiates a poison and a remedy.” Even water, when consumed too quickly and in inordinate amounts, can be lethal.

One of the primary objectives of medical geologists, therefore, is to determine the optimal exposures for people to the essential elements in order to maintain or improve health. Diseases such as hyperkalemia (due to excess potassium), hypercalcemia (due to excess calcium), and hyperphosphatemia (due to excess phosphorus) exist in various parts of the world even though potassium, calcium, and phosphorus are essential dietary elements.

I have studied a number of specific cases that illustrate how optimal amounts of trace elements are needed to maintain good health and how an imbalance in these elements can lead to disease.

This yearlong essay series celebrates 125 years of *Science* by inviting researchers from around the world to provide a regional view of the scientific enterprise. Series editor, Ivan Amato

Chandra Dissanayake Sri Lanka



Chandra Dissanayake, a senior professor of geology at the University of Peradeniya, Sri Lanka, has pioneered geochemical research in Sri Lanka. He is the only Sri Lankan to have obtained both D.Phil. and D.Sc. degrees from Oxford University, UK. From the very beginning of his research career, he sought avenues off the beaten track of routine geology by exploring the boundaries between geology and other scientific areas. For nearly 30 years he has carried out research in the little-known field of medical geology and has published many research papers on the topic. He is a recipient of the National Award for Science in Sri Lanka and the Gold medal of the Institute of Chemistry, Sri Lanka, in recognition of his groundbreaking research on applied geochemistry. He is a fellow of the Academy of Sciences for the Developing World (TWAS), headquartered in Trieste, Italy, and an Alexander von Humboldt Fellow of Germany. In addition to his scientific pursuits, he is engaged in social work and is the president elect of the Lions club of Kandy, Sri Lanka. Recently he was appointed as the chairman of the Sri Lanka Pugwash Council, whose members seek peaceful and cooperative solutions to armed conflicts and other global problems.

All essays appearing and interactive features in this series can be found online at www.sciencemag.org/sciext/globalvoices/

Fluoride in Drinking Water—Friend and Foe

The link between the fluoride geochemistry of water in an area and the incidence of dental fluorosis, a tooth-damaging condition, is a well-established relationship in medical geology. Although the value of limited exposures to fluoride for human health, particularly for dental health, is well known, higher exposures in many tropical lands where fluoride is found in excessive quantities in the drinking water have had devastating effects. As in the case of some essential trace elements, the optimal level of fluoride varies within a narrow range. In locations where the exposure to fluoride is not well controlled, including my home country, Sri Lanka, many people suffer from a fluoride imbalance. While carrying out detailed research on the medical geology of dental fluorosis in Sri Lanka, I have seen firsthand how disfiguring this condition can be.

When the fluoride content of the drinking water exceeds about 1.5 mg/liter, the maximum concentration recommended by the World Health Organization, and when such fluoride-rich water is consumed, particularly by children under 7 years of age, the teeth develop a dark brown coloration and a mottling. This condition is known as dental fluorosis (see the top figure). Although not a life-threatening condition, it constitutes a massive social problem, particularly for girls worried about their marriage prospects. Some use sand paper to brush their teeth with the hope of getting rid of the ugly stains.

Sri Lanka, an equatorial developing country with a population of about 20 million, has well-defined dry and wet zones. In the dry zone, dental fluorosis is highly prevalent, and a population of more than 2 million is at risk of developing it.

Most Sri Lankans live in close association with their immediate geological environment, and only about 30% have clean piped water with controlled mineral content. The rest generally get their drinking water from wells. In some dug wells, and most notably in deep boreholes, the fluoride concentration in water exceeds 1.5 mg/liter. In some cases, the concentration can be as high as 10 mg/liter. The sources of the fluoride are the high-grade metamorphic rocks in the dry zone of Sri Lanka. These rocks include an abundance of fluoride-bearing minerals such as mica, hornblende, and fluorite. With funding from the Natural Resources Energy and Science Authority of Sri Lanka, I was able to produce a map for Sri Lanka showing the fluoride-rich zones and their potential impact on dental health (see the bottom figure).

From a strictly scientific perspective, one of the most interesting aspects of these studies is the biomineralogy of tooth enamel and



Dental danger. Geochemists studying the quality of water in a well in the dry zone of Sri Lanka. (Inset) Teeth from a patient with dental fluorosis in Sri Lanka.

the process by which hydroxyapatite, the primary mineral in teeth and bones, transforms into fluoroapatite when fluoride ingestion is excessive.

The "Geochemical Disease": Iodine Deficiency

It has been estimated that nearly 30% of the world's population is at risk for some form of iodine deficiency disorder (IDD). Insufficient intake of iodine is the world's most common cause of mental retardation and brain damage with 1.6 billion people at risk, 50 million children already affected, and 100,000 more adding to their ranks every year.

IDDs are particularly severe in tropical regions. The resulting large populations of people with impaired mental function have serious direct and indirect impacts on all aspects of life in these places.

The geochemistry of iodine and its chemical species has a marked influence on the prevalence of IDD, including endemic goiter (see figure on the next page). Collectively, these IDD are often referred to as "geochemical diseases" in view of their etiology in the geological environment. The sea is a major source of iodine, so there often is a relationship between the incidence of IDD in a region and that region's distance from the sea. In general, the farther away from the sea, the less iodine is available. Other factors such as atmospheric circulation, however, may play a role in iodine availability, as does topography. In many mountainous regions, for example, iodine abundance is quite low, with a concomitant increase in IDD.

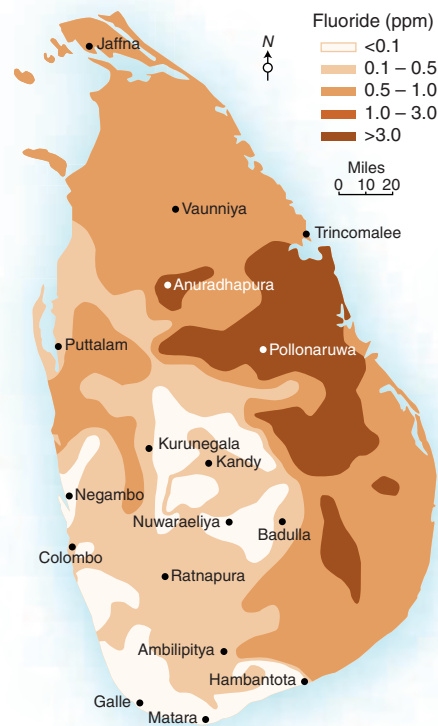
Humic substances rich in organic matter in the environment also are known to play a major role in the speciation and geochemical mobility of chemical elements such as iodine. The in-ground conversion of chemical species into toxic or nontoxic forms has important implications for the health of individuals living in a particular geochemi-

cal habitat. Iodine is strongly fixed by humus, and soils rich in humus therefore tend to be rich in iodine. However, the bioavailability of that iodine, which depends on the pH conditions, often is very low. Because iodine is strongly fixed by clay and humus, these materials may function as "geochemical goitrogens," particularly in the tropical environment. Interestingly, in the Kalutara district—an area close to the sea on the west coast of Sri Lanka (and there-

fore only minimally affected by last December's tsunami)—the endemic goiter rate is high despite the proximity to the sea. This clearly supports the hypothesis that the environment contains goitrogens, which trap the iodine and make it far less available to the people living in the area.

Eating Clay

The habit of eating soil and the physiological effects of its ingestion are a fascinating field of study. This seemingly repulsive phenomenon enables researchers to study the direct link between soil chemistry, soil mineralogy, and health.



Map showing fluoride distribution in Sri Lanka.

CREDIT: (TOP AND BOTTOM) CHANDRA DISSANAYAKE

Geophagy is defined as the deliberate and regular consumption of earthy materials such as soils, clays, and mineral substances by humans and animals, among them elephants, monkeys, chimpanzees, gorillas, birds, reptiles, and horses. In some cases, animals eat pebbles and rocks that serve as mechanical aids to digestion, for grinding food into bits.

The practice of geophagy by humans has been observed on all continents, although it is most commonly seen in the tropics, and particularly in tropical Africa. It is especially common among pregnant women.

Alexander von Humboldt, who explored South America for its natural resources, observed the practice of geophagy during his expeditions to Orinoco in Venezuela in the period from 1799 to 1804. The Ottomac people, who practiced geophagy there, apparently did not eat every type of clay, but chose only those clays that were most “unctuous and smoothest to touch.” Interestingly, the Ottomacs did not suffer health problems as a result of their clay eating, yet other tribes who ate different soils did become sick.

This early observation by von Humboldt generated considerable interest among medical scientists years later. The debate about the possible benefits of eating soil continues today.

Why do humans and animals consume soil? Could it be that inorganic nutrients in the soil supplement our dietary intake of essential trace elements? Does the ingestion of soil cause detoxification of noxious or unpalatable compounds present in the diet? Do these soil elements alleviate gastrointestinal ailments? These questions need to be answered, and further detailed research will be needed to understand the strange phenomenon of geophagy.

Is Hard Water Good for Heart Ailments?

One of the most tantalizing geology-health correlations involves the incidence of cardiovascular diseases (CVDs) and the water hardness of a particular area. In several countries and regions, a negative correlation between water hardness and deaths due to CVD has been observed. This correlation has been seen in both temperate and tropical countries. Even though a causal effect still cannot be ascribed to this geochemical correlation, the potential role that trace elements in drinking water could play in this relationship has aroused considerable curiosity among medical geologists.

If we accept for now that there is some causal basis to this correlation, then the question to ask is this: What is it in the hard water that is cardio-protective?

Mounting evidence from many studies indicates that this “water factor” is magne-

sium, with calcium playing a supportive role. The presence of calcium and magnesium in natural water results from the decomposition of calcium and magnesium aluminosilicates, which derive from limestone, magnesium limestone, magnesite, gypsum, and other minerals.

An important point to note is that only two out of every three studies on this topic have shown a correlation between cardiovascular mortality and water hardness. Studies probing the effect of water magnesium alone have all shown an inverse correlation between cardiovascular mortality and water magnesium level—the more magnesium, the lower the rate of CVD mortality.

Even though medical geologists have shown much enthusiasm for the possible cardio-protective role of magnesium, those in the medical profession are yet to be fully convinced of the hard water–CVD connection. More research is needed to clearly pinpoint the elusive “water factor,” if indeed there is one to be found.

The Radiation Paradox

Natural radioactivity on Earth has been in existence since the planet formed, and there are about 60 radionuclides

present in nature. These are found in air, water, soil, rocks and minerals, and food. About 82% of this environmental radiation is from natural sources, the largest of which is radon.

Some areas of the world, called high background radiation areas (HBRAs), have anomalously high levels of background radiation. In such terrains, the geology and geochemistry of the rocks and minerals have the greatest influence in determining where the high natural radiation shows up. Extreme HBRAs are found in Guarapari (Brazil), southwest France, Ramsar (Iran), parts of China, and the Kerala coast (India). Of these, most are found in tropical, arid, and semiarid areas. In certain beaches in Brazil, monazite sand deposits are abundant. The external radiation levels on these black beach sands range up to 5 mrad/hour, which is nearly 400 times the normal background level in the United States. The Brazilian coastal sands have several radioactive minerals, among them monazite, zircon, thorium, and niobate-tantalate, as well as nonradioactive minerals, including ilmenite, rutile, pyrochlore, and cassiterite.

In India, along the 570-km-long coastline of Kerala, there are major deposits of monazite-rich mineral sands with very high natural radiation. The monazite deposits are larger than those in Brazil, and the dose from external radiation is, on average, similar to those reported in Brazil.

Ramsar, a city in northern Iran, has one of the highest natural-radiation levels in the world. In some locations at Ramsar, the radiation level is 55 to 200 times higher than the background level. Exposures as high as 260 mGy/year have been recorded in Ramsar. The unit of ionizing radiation here, grays per year, corresponds to 1 J of energy imparted to 1 kg of tissue (the milligray, mGy, which is one-thousandth of a gray, is more commonly used). Whole-body exposure to a uniform dose of 3 to 5 Gy would kill 50% of those exposed within 1 or 2 months.

The most interesting feature in all these cases is that the people living in these HBRAs do not appear to suffer any adverse health effects as a result of their high exposures to radiation. On the contrary, in some cases the individuals living in these HBRAs appear to be even healthier and to live longer

than those living in control areas that are not classified as HBRAs. These phenomena pose many intriguing questions for medical geologists.

Breaking Barriers

The examples that I have outlined here illustrate that geology is no longer confined to

the study of rocks and minerals. I suspect that as the field of medical geology becomes better known, the medical community will discover that geology could play a major role in the etiology of a variety of diseases. Uncovering these relationships is an inherently multidisciplinary task. After all, for medical geology and geochemistry to be of use to the public and to health authorities of a given region, all samples from the local environment—that is, rock, soils, water, plants, and food—need to be studied together and correlated with *in vivo* studies.

In the medical field, too, as discoveries are made in the physiology and metabolism of trace elements, the biochemical mechanisms underlying the body's absorption and rejection of trace elements, and the cellular mechanisms that regulate these processes, researchers may also find a greater need for the understanding of geology and trace-element geochemistry.

Geology and medical science, disciplines that until now have been considered poles apart, may now find themselves joined into a multidisciplinary framework for unraveling some of nature's most interesting secrets.



Woman with endemic goiter.

The author is at the University of Peradeniya, Peradeniya, 20400 Sri Lanka. E-mail: cbdissa@hotmail.com

10.1126/science.1115174

Suppressing Cancer: The Importance of Being Senescent

Judith Campisi

Cancer is a potentially lethal disease in mammals and other complex organisms with renewable tissues. Tumors originate from cells that are actively dividing. Such cells are at much greater risk than postmitotic (nondividing) cells for acquiring mutations, a major driving force for cancer development. Cell division is extensive during development and continues during maturation and adulthood. Yet cancer is typically an age-related disease, developing primarily in older adults. Why, then, don't mammals develop cancer earlier and more frequently? The answer lies in the tumor suppressor mechanisms that evolved to protect complex organisms from malignant tumors (1). Some of these mechanisms protect the genome from damage or mutation. Others eliminate or arrest the proliferation of potential cancer cells by processes called apoptosis or cellular senescence. There is ample evidence that apoptosis, or cellular suicide, suppresses tumorigenesis *in vivo*. However, evidence that cellular senescence, the permanent arrest of cell division, suppresses cancer has been largely circumstantial. Four recent papers dispel doubts that cellular senescence is an important anticancer defense *in vivo* (2–5). Furthermore, they show that activated oncogenes—mutant genes that have the potential to transform normal cells into a cancerous state—induce cellular senescence *in vivo* (see the figure), a phenomenon that previously had been seen only in cell culture. The findings support the idea that the senescence response is a failsafe mechanism that prevents the proliferation of cells at risk for neoplastic transformation.

Cellular senescence was first identified as a process that limits the ability of normal human cells to proliferate in culture. We now know that this limit is caused by at least two intertwined mechanisms (1). First, the erosion of telomeres, regions at the ends of chromosomes that stabilize DNA, elicits

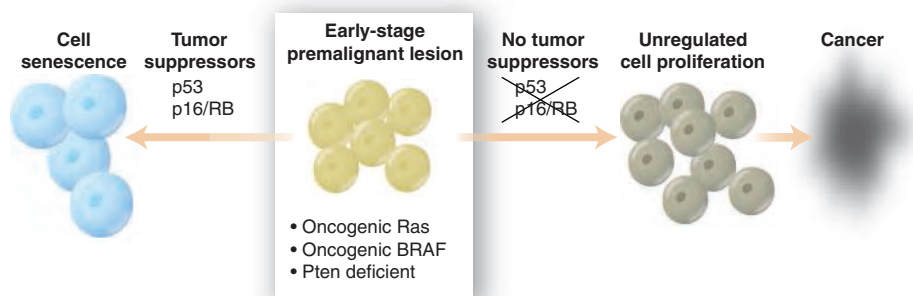
a DNA damage response that causes the cell division cycle to arrest. This response requires the activity of a signaling pathway that includes the tumor suppressor protein p53. Second, cumulative stress of an unknown nature induces expression of the p16 tumor suppressor. This activates a signaling pathway that involves another tumor suppressor protein, pRB, which in turn halts cell cycle progression. We also now know that many stimuli induce a senescence response. These include nontelomeric DNA damage that engages the p53 pathway and certain oncogenes that trigger the p16-pRB pathway. How relevant is the senescence response, particularly the response to oncogenes, *in vivo*?

Collado *et al.* (2) used a mouse model in which inducible expression of an oncogenic allele of the Ras gene (*K-ras*^{V12})

typically developed were largely devoid of senescent cells, it is likely that rare variants overcome senescence and then progress to full-blown malignancy.

Using an entirely different mouse model in which the Pten tumor suppressor is inactivated in the prostate at puberty, Chen *et al.* (3) similarly found that premalignant or nonlethal cancers expressed senescence markers, whereas malignant tumors did not. Pten dampens growth-promoting signals. Its absence in mice resulted in lethal invasive prostate cancer only when p53 was inactivated. In culture, Pten-deficient cells from these animals entered a senescent state that was overcome by loss of p53 function. Thus, p53 restricts the growth and malignant progression of Pten-deficient cells—both in culture and *in vivo*—by inducing cellular senescence. This is very likely relevant to human prostate cancer because cells expressing a senescence marker were found in early-stage human prostate cancers, but not in frankly malignant tumors.

Michaloglou *et al.* (4) explored the role of cellular senescence in human cancer progression by examining benign melanocytic tumors (nevi), many of which express an oncogenic form of BRAF, a downstream



Cell senescence as an anticancer defense. Mutations that activate the oncogenes Ras or BRAF or inactivate the tumor suppressor Pten produce premalignant (early-stage) lesions. If there is additional loss of p53 or p16/pRB function, premalignant cells cannot undergo cellular senescence and progress instead into malignant tumors.

causes multiple lung adenomas, a few of which progress to malignant adenocarcinomas. Ras proteins transduce growth factor signals and oncogenic Ras forms deliver unregulated signals. Collado *et al.* show that the adenomas, but not adenocarcinomas, express several markers characteristic of senescent cells in culture. Likewise, they observed premalignant lesions in the pancreas and skin that expressed senescence markers, whereas malignant tumors did not. Thus, activated oncogenes induce a senescence response *in vivo*, at least in mice. Because malignant tumors that even-

mediator of RAS. When expressed in cultured human melanocytes and fibroblasts, mutant BRAF caused a transient burst of proliferation, followed by senescence and increased p16 expression. This arrest was overcome by expressing the viral oncogene SV40 T antigen, which inactivates both p53 and pRB. In human skin samples, the melanocytes present in nevi expressed markers of senescence. It may be that these lesions are benign because of senescent melanocytes that harbor BRAF mutations. Interestingly, p16 expression in nevi was heterogeneous and its elimination had no

Lawrence Berkeley National Laboratory, Life Sciences Division, 1 Cyclotron Road, Berkeley, CA 94720 and Buck Institute for Age Research, 8001 Redwood Boulevard, Novato, CA 94945, USA. E-mail: jcampisi@lbl.gov

effect on senescence. Moreover, telomere erosion was not apparent, suggesting that some cells senesce by telomere- and p16-independent mechanisms. One possible mechanism is DNA damage caused by the reactive oxygen species that mediate RAS-dependent mitogenic signals (6).

Finally, Braig *et al.* (5) used a mouse model in which oncogenic Ras (Eμ-*N-Ras*) is constitutively expressed in hematopoietic cells. The study shows that a deficiency in Suv39h1, a histone methyltransferase, markedly accelerates the development of lethal tumors. Suv39h1 is thought to promote the heterochromatic silencing of growth-promoting genes in senescent cells. This silencing causes the senescence response of lymphocytes to oncogenic Ras. Lymphomas that develop in Eμ-*N-Ras* mice

undergo senescence in response to chemotherapy, but this did not occur in Suv39h1-deficient tumors. Rather, Suv39h1-deficient tumor cells underwent apoptosis. Thus, cell senescence suppressed lymphomagenesis in these mice.

Together, these papers support the idea that cellular senescence, like apoptosis, plays an important role in suppressing tumorigenesis in mice and humans in vivo. Needless to say, many questions remain. What are the mechanisms that determine whether cells undergo senescence or apoptosis when challenged by potentially oncogenic insults? Are there pathways other than the p53 and p16-pRB pathways that cause the senescence response? And, is senescence as effective as apoptosis at preventing cancer? The latter question is especially

important because senescent cells secrete factors that can stimulate the proliferation and malignant progression of neighboring cells (7, 8). And thus, a potential irony lurks: Prolonged presence of senescent cells may eventually facilitate the development of malignant cancers from benign lesions.

References

1. J. Campisi, *Cell* **120**, 513 (2005).
2. M. Collado *et al.*, *Nature* **436**, 642 (2005).
3. Z. Chen *et al.*, *Nature* **436**, 725 (2005).
4. C. Michaloglou *et al.*, *Nature* **436**, 720 (2005).
5. M. Braig *et al.*, *Nature* **436**, 660 (2005).
6. K. Irani *et al.*, *Science* **275**, 1649 (1997).
7. A. Krtolica, S. Parrinello, S. Lockett, P. Desprez, J. Campisi, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12072 (2001).
8. B. D. Chang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 389 (2002).

10.1126/science.1116801

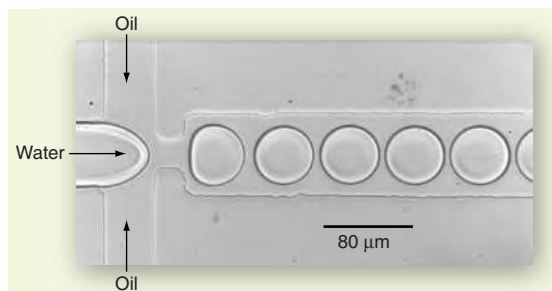
APPLIED PHYSICS

Droplet Control for Microfluidics

Mathieu Joanicot and Armand Ajdari

Nanoliter droplets of uniform size spontaneously form in microchannels when two immiscible fluid streams merge (1) (see the figure). This nonlinear process involves basic physics (2), with the local geometry and surface chemistry of the microchannel strongly affecting the competition between viscous forces tending to draw the fluids along the channel and capillary forces tending to form droplets so as to minimize the total interface between the two fluids. This results in droplets of sizes comparable to the channel diameter (3), and these sizes can be tuned by adjusting the flow rates of the various streams (4, 5). The determination of which liquid is inside the droplet and the extension of the regime where droplets are emitted periodically are controlled by the wetting competition (which fluid preferentially interacts with the channel surfaces) and by added surfactants (6). This new route for easy and steady production of calibrated emulsions opens a stimulating field for applications of microfluidic devices [for a recent review, see (7)].

Typical microfluidic channel sizes (height and width) are in the range of 10 to 100 μm, and flow rates are between 10 and 1000 nL/s. This leads to nanoliter-size droplets, produced at frequencies of 10 Hz to 10 kHz, moving at speeds from microm-



Tunable droplets. Water droplets form at a rate of 1000 per second in an asymmetric microchannel containing hexadecane. The channel is 30 μm high.

eters per second to centimeters per second. As a result, these tiny droplets are almost ideal chemical reactors because they create homogeneous controlled conditions (8). In the first place, the very high surface-to-volume ratio (owing to the small size of the droplets) grants very fast thermal transfer. In addition, each droplet moves as an independent nanoliter batch reactor, with no hydrodynamic dispersion. For a steady flow, each location along the channel directly corresponds to a unique residence time after droplet formation. And internal recirculation within the droplet permits fast and efficient mixing, especially if wiggly channels are used (8). Such ideal reactors allow one to follow reactions in time. An interesting tool for such studies is confocal Raman microspectroscopy, which can be used to determine the chemical composition anywhere along the channel, providing a mapping that is a direct measurement of the whole kinetics at once. For short dis-

tances and high flow rates, reaction times as short as a few milliseconds can be measured (8). By varying the initial composition of the droplets, one can assess its effect on yield and kinetics, providing a useful oper-

ational research tool for laboratories. Control of residence time also enables synthesis of quite monodisperse small particles within the droplets.

There is another area in which these devices can be valuable tools for materials engineering, namely the transformation of each droplet into a single colloidal object. For example, one can dissolve the desired molecules or polymers into an organic phase and flow the latter into an aqueous stream to generate droplets. To dry the resulting emulsion, the organic solvent is either exchanged with the aqueous phase or slowly evaporated through it. A last step of ultraviolet-induced cross-linking or polymerization can then be used to solidify the colloids. The polydispersity of the particles can be as low as a few percent, far better than what is achievable with classical means of generating emulsions. Colloids of various shapes (disks, cylinders, and so forth) can be obtained by solidifying confined droplets (9). Monodisperse droplets of liquid crystals can also be obtained (10).

As suggested earlier, the use of hydrophobic channels results in formation of water droplet in a stream of hydrophobic liquid (oil), whereas hydrophilic channels favor creation of water-in-oil droplets. Now imagine a two-step process whereby a hydrophobic channel, in which a water-in-oil emulsion is generated, connects to a hydrophilic channel in which water flows. This can result in a multiple emulsion of

water-in-oil-in-water (11) by encapsulation of the initial simple emulsion. In principle, the size, the number, and even the composition of the drops can be controlled at each step of the process, leading to a unique way of producing monodisperse, perfectly controlled, multiple emulsions. However, mastering the surface properties of the channel is a prerequisite for this process to run smoothly and controllably. A way around this is to produce the multiple emulsion in one step, using three concentric flows so as to avoid contact of the inner streams with the channel wall. Such a geometry is difficult to fabricate within the usual two-dimensional microfluidic structures obtained by lithography, but it was recently achieved with a clever multiple-micropipette device. Monodisperse core-shell particles and capsules (12) were obtained, as well as polymersomes of controlled sizes synthesized from block copolymers (13).

Microfluidics can thus be a powerful and versatile tool for materials or colloidal engineering. It is the only technique that can produce 100% encapsulation of an active substance by means of a one-step process. Engineering new controlled materials then becomes a game limited only by the number of combinations of the basic products, and by one's imagination.

A controversial question naturally follows: Besides being a smart research tool, can microfluidic droplet devices be used as production tools (chemical plants on a chip)? With a flow rate of 1000 $\mu\text{L}/\text{hour}$ per channel, one needs 1000 channels in parallel to produce a liter of material per hour. High levels of integration can be reached on a chip, but likely a linking up of many devices will be necessary, potentially leading to complications for connections and control. Developing the corresponding technology may be economically sensible only for materials with very high added value (such as those in biology, pharmaceuticals, or cosmetics), but not for conventional chemistry or basic material production.

A large number of research reports and patents attest to the ongoing attempts to develop new research tools for biological or pharmaceutical applications by means of droplet microfluidics. The goal is to handle the many droplets that can be generated with only a minute amount of material, and to divide and recombine them in a multiplicity of nanoreactors so as to perform high-throughput screening and combinatorial studies (14). Some would like to see this approach as the next-generation technology that will replace the widely used combinatorial robotic platforms.

What is required is on-chip control and reproducibility of many droplet processes: fabrication, sorting, storage, fusion,

breakup, and trafficking (3, 8, 15), among others. For example, droplet generation is currently controlled by tuning the input flow rates. Unfortunately, this affects simultaneously the frequency, size, composition, and speed of the droplets, whereas one would want to control each of these parameters independently. A natural microfabrication strategy is to integrate actuators to achieve local control of droplet motion. Pneumatic actuators have proven highly integrable in a different microfluidic context (16), and electrostatic actuation with integrated electrodes is also being investigated. The gain in control may unfortunately result in somewhat sophisticated and specialized chips with limited flexibility and versatility. It may be necessary to standardize a few basic on-chip functions, with a drift toward passive strategies that often combine simplicity and robustness.

Improvements in design should be guided by attention to basic physics and chemistry. Interfaces are essential at such small scales, and surfactants should be avoided as much as possible. Hence, a synergy between microfabrication requirements and surface chemistry is needed to yield robust channel wetting properties. This will create the needed reproducibility of droplet generation for a range of flowing liquids and flow rates. The complex dynamics of thousands of droplets flowing in channels with long-range hydrodynamic

correlations will require modeling, possibly with the help of concepts borrowed from dynamical system theories. The result will be the design of smart network topologies and (ideally passive) functionalities (15), opening the way to control, reproducibility, and versatility in on-chip droplet management, ultimately at the droplet level.

References

1. T. Thorsen, R. W. Roberts, F. H. Arnold, S. R. Quake, *Phys. Rev. Lett.* **86**, 4163 (2001).
2. H. Stone, A. D. Stroock, A. Ajdari, *Annu. Rev. Fluid Mech.* **36**, 381 (2004).
3. D. R. Link, S. L. Anna, D. A. Weitz, H. A. Stone, *Phys. Rev. Lett.* **92**, 054503 (2004).
4. J. D. Tice, A. D. Lyon, R. Ismagilov, *Anal. Chim. Acta* **507**, 73 (2004).
5. P. Guillot, A. Colin, in preparation.
6. R. Dreyfus, P. Tabeling, H. Willaime, *Phys. Rev. Lett.* **90**, 144505 (2003).
7. K. Jensen, A. Lee, *Lab. Chip* **4**, 432N (2004).
8. H. Song, J. D. Tice, R. F. Ismagilov, *Angew. Chem. Int. Ed.* **42**, 768 (2003).
9. D. Tsoi, K. Hattori, T. A. Doyle, *Langmuir* **21**, 2113 (2005).
10. A. F. Nieves *et al.*, *Adv. Mater.* **17**, 680 (2005).
11. S. Okushima, T. Nisikawa, T. Torii, T. Higuchi, *Langmuir* **20**, 9905 (2004).
12. A. S. Utada *et al.*, *Science* **308**, 537 (2005).
13. E. Lorenceau *et al.*, *Langmuir*, in press.
14. B. Zheng, L. S. Roach, R. F. Ismagilov, *J. Am. Chem. Soc.* **125**, 11170 (2003).
15. Y. Tan, J. S. Fisher, A. I. Lee, V. Cristini, A. Phillip, *Lab. Chip* **4**, 292 (2004).
16. T. Thorsen, S. J. Maerkl, S. R. Quake, *Science* **298**, 580 (2002); published online 26 September 2002 (10.1126/science.1076996).

10.1126/science.1112615

PLANETARY SCIENCE

The Enigma of the Martian Soil

Amos Banin

Since the two Viking landers touched down on Mars in 1976, the ubiquitous surface soil and dust have defied attempts to model their properties and understand how they formed. The fine soil that gives

Mars its red color holds clues to the presence of liquid water, the rock

weathering processes, and the potential biological history of the planet. Through dust storms, it plays a key role in the climatic cycles of Mars. It may hamper exploration by interfering with robotic and human performance, but may also offer a valuable resource by supplying water and fuel and supporting plant

growth and food production. What is known and what is still puzzling about the martian soil? And do we have similar soils on Earth? Recent results from the Mars exploration rovers Spirit and Opportunity and from terrestrial studies add to our understanding and open new questions.

Data from the Viking landers and orbiters and the Pathfinder rover (1–6) show that the surface of Mars is covered by a blanket of fine-textured soil compositionally similar to the atmospheric dust (see the figure). The soil contains silicon, iron, aluminum, magnesium, calcium, titanium, sulfur, and chlorine at unique elemental proportions (that is, relatively rich in sulfur and chlorine compared to most terrestrial soils). It lacks organic matter and shows strong oxidizing activity—a combination that suggests the possibility of rapid chemical decomposition of organic matter leading to the decimation of living organisms. It is rich in

The author is in the Department of Soil and Water Sciences, Hebrew University, Rehovot 76100, Israel, and at the SETI Institute, Mountain View, CA 94043, USA. E-mail: amos.banin@huji.ac.il

amorphous weathered silicate and iron oxide grains but lacks grains of well-developed secondary silicates such as clays. It also lacks carbonates but contains high concentrations of chlorine and sulfur salt-minerals (evaporites) and magnetically active minerals.

More recent results obtained by the Mars exploration rovers Spirit and Opportunity provide evidence for localized formations of water-modified weathered rocks that contain the mineral jarosite (7–10). This mineral typically forms in aqueous environments of highly acidic sulfate solutions. On Earth, it is found near mine spills and in acid sulfate soils. In general, however, the elemental composition of the soils analyzed by the two rovers is similar to that summarized above.

Using onboard spectroscopic tools, the rovers established the pervasive presence of nanometer-scale iron oxide particles in the soils. The use of the rover wheels for trenching and of a brush and a rock abrasion tool for rock cleaning enabled the acquisition of depth profiles of elemental concentrations and mineral types in the soils and on the surface of rocks at the rover sites. The results show that the nanometer-scale iron oxide particles and the evaporites are usually more concentrated at the soil-atmosphere and rock-atmosphere interfaces (7, 10–12).

This evidence poses a number of questions (13). Where is all the organic matter imported by meteorites or produced otherwise? Is the soil really highly oxidizing? If so, to what extent and as a result of what formation mechanism? How are the soluble evaporites emplaced in the soil? Are they really water-transported evaporites? Do they provide evidence for extensive water on the martian surface? If yes, then where are the other minerals, such as clays and carbonates, that should form during low-temperature hydrothermal weathering? Why are the abundant secondary iron oxides and silicates in the soil primarily in the form of nanometer-scale crystals rather than larger, more mineralogically stable grains? Is the soil coupled chemically and mineralogically to the local surface rocks? And is there any connection to extinct or extant life in the soil?

Several hypotheses have been proposed to explain the formation of the soil and answer these questions. The “acid fog” scenario (13, 14) suggests that a major portion of the soil formed by weathering at the microscopic rock–atmosphere interface over hundreds of millions of years. This unique weathering process is driven by mists and fogs of acidic volcanic aerosols, which are neutralized through chemical reactions with mineral components of the surface

rocks. No liquid water is involved in this weathering process, thus preventing the growth of larger crystals and the formation of secondary minerals such as clays. In other hypotheses, liquid water—in the form of either shallow seas (15) or seeping-up acidic groundwater (16)—creates precipitation environments that are conducive to the formation of jarosite, for example.

There is some debate over whether Mars is really covered with soil. Squyres *et al.* define martian soil as “any loose, unconsolidated materials that can be distinguished from rocks, bedrock, or strongly cohesive sediments. No implication of the presence or absence of organic materials or living matter is intended” [(8), p. 799; (9), p. 1702]. I also favor the use of the term soil rather than the term “regolith” (defined as a layer of loose heterogeneous material covering solid planetary rock) or the often-used term “dirt” to describe martian soil. However, the above definition is perhaps too narrow and can be broadened to agree more precisely with the general definition of soil on Earth.

Terrestrial soils may be defined as the top weathered or partly weathered layer of the terrestrial lithosphere that is exposed to and modified by atmospheric, hydrolytic, and biotic effects. Soil consists of a porous mineral matrix that contains a mixture of weathered and non-weathered rock grains, organic matter, and debris of dead organisms. The weathering agents are physical (such as thermal expansion/contraction and freeze/thaw cycles) and chemical (mostly hydrolytic processes but also reduction/oxidation and other reactions).

With current evidence, martian soils can be defined, analogously to terrestrial soils, as the top nonconsolidated layer of weathered and/or partly weathered rocks of the martian lithosphere that is exposed to and modified by atmospheric and hydrolytic effects. Note that this definition omits the biotic effect. It cannot be ruled out that part of the soil formed early in Mars history, when conditions on its surface may have been much like those on the primi-



Viking 2 Lander sampling-arm trench and frost patches



Pathfinder-Sojjourner wheel marks



Mars exploration rover Spirit surface scraping



Mars exploration rover Opportunity touchdown prints

Back to the soil. Close-ups of Mars soil taken by landers and rovers.

tive Earth, so that abiotic and primitive biotic evolution could have taken place and rock weathering that formed soils was enhanced. However, current evidence suggests that rock weathering and soil formation on Mars is a very slow process that continues to the present. These weathering processes are driven by the chemical interactions of gaseous and aqueous species, transported by the martian atmosphere and hydrosphere, with the martian rocks. For example, the typical red color of Mars is due to the oxidation of iron (“rusting”) in exposed rock minerals by oxidizing gases that form photochemically in the atmosphere. The sulfate and chloride salt evaporites may have been added to the soil as it formed by acid fog weathering caused by acidic volcanic volatiles transported in the atmosphere. The martian hydrosphere influenced soil formation through outflows of water (either fresh water or highly saline brines), catastrophic floods, and/or the melting of

permafrost by shallow volcanism. The released water interacted with martian rocks and modified their mineralogy through dissolution and reprecipitation.

The chemical interactions with the atmosphere and hydrosphere profoundly altered the mineralogy and the reactivity of the top layer of the martian lithosphere, rendering it much more chemically reactive than the original lava and forming the global martian soil layer (pedosphere). Martian soil thus has several attributes that render it a soil rather than a regolith. In contrast, the Moon is covered with a regolith composed of primary rock particles broken down physically or mechanically, mostly by micrometeorite impacts, but never exposed to atmospheric effects or hydrolytic weathering.

Are there terrestrial sites, soils, or mineral mixtures that are true analogs for martian soil? Navarro-González *et al.* (17) have argued that soils in one of the driest regions on Earth, the Atacama Desert in Chile, are a model for martian soils. They found that the number of culturable bacteria decreased steeply along a

north-south transect, reaching extremely low values at the driest site. Using simulations of the Viking biology experiments (3, 4), they showed that pyrolyzable organic matter decreased to extremely low values along the transect from the wetter to the drier sites; abiotic formate decomposition was measured in all the sites. In this respect, Atacama soils may indeed be the closest model for Mars soils on Earth.

However, the findings are controversial, and many questions remain open. Maier *et al.* (18) have reported orders-of-magnitude higher counts of culturable bacterial cells in samples from the same sites but taken from deeper horizons in the soil profile, not just from the top surface layer as reported in (17). Furthermore, detailed calculations and additional analyses suggest that the Atacama soils contain far higher concentrations of total organic carbon than the Mars soils analyzed in the Viking pyrolysis experiment may have had, considering the system's detection limits for organic carbon [in the range of a few parts per billion (ppb) for most compounds and up to several hundred ppb for a few compounds]. Recalculation of the data reported in figure 2C of (17) shows that the total organic carbon released as

benzene and formic acid from the Atacama soils varied from ~13,000 ppb in the driest sites to ~85,000 ppb in the wettest sites—much higher than the detection limits of the Viking pyrolysis instruments (3).

Is the martian soil completely devoid of organic matter, as concluded by the Viking team? Or was the protocol used by the Viking pyrolysis instrument limited in its detection ability (19)? The answer to this question will profoundly affect our view of the presence or absence of life on Mars. Future missions to Mars should reanalyze, as a high priority, the total concentrations of organic carbon and the biologically important element nitrogen in the soils. This will contribute to our understanding of the soil's mode of formation, its reactivity and toxicity, utility as a resource, and astrobiological connections.

The above questions notwithstanding, the Atacama Desert is a unique spot on Earth. It behooves the scientific community to carefully control and preserve sections of this environment for future study. Selected sites in this desert could be declared a "nature preserve" to prevent the contamination of this ecosystem by excessive human activity. Thus pre-

served, these sites could continue to be used as a testing ground for sterilization and quarantine protocols developed for vehicles and missions sent to Mars and for careful field experimentation with samples returned from Mars by future missions.

References and Notes

1. P. Toulmin III *et al.*, *J. Geophys. Res.* **82**, 4625 (1977).
2. B. C. Clark *et al.*, *J. Geophys. Res.* **87**, 10059 (1982).
3. K. Biemann *et al.*, *J. Geophys. Res.* **30**, 4641 (1977).
4. G. V. Levin, P. A. Straat, *J. Geophys. Res.* **82**, 4663 (1977).
5. A. Banin, B. C. Clark, H. Wänke, in *Mars*, H. H. Kieffer, B. M. Jakosky, C. W. Snyder, M. S. Matthews, Eds. (Univ. of Arizona Press, Tucson, AZ, 1992), pp. 594–625.
6. R. Rieder *et al.*, *Science* **278**, 1771 (1997).
7. G. Klingelhöfer *et al.*, *Science* **306**, 1740 (2004).
8. S. W. Squyres *et al.*, *Science* **305**, 794 (2004).
9. S. W. Squyres *et al.*, *Science* **306**, 1698 (2004).
10. R. V. Morris *et al.*, *Science* **305**, 833 (2004).
11. R. Gellert *et al.*, *Science* **305**, 829 (2004).
12. R. Rieder *et al.*, *Science* **306**, 1746 (2004).
13. A. Banin *et al.*, *J. Geophys. Res.* **102**, 13341 (1997).
14. N. J. Tosca *et al.*, *J. Geophys. Res.* **109**, E05003 (2004).
15. S. W. Squyres *et al.*, *Science* **306**, 1709 (2004).
16. R. G. Burns, D. S. Fisher, *J. Geophys. Res.* **95**, 14415 (1990).
17. R. Navarro-González *et al.*, *Science* **302**, 1018 (2003).
18. R. M. Maier *et al.*, *Science* **306**, 1289 (2004).
19. S. A. Benner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2425 (2000).
20. Supported in part by NASA Astrobiology Institute/SETI Institute Astrobiology Project NNA04CC05A.

10.1126/science.1112794

DEVELOPMENTAL BIOLOGY

Rac1 Up for Epidermal Stem Cells

G. Paolo Dotto and George Cotsarelis

A well-known dermatologist once noted, "No one ever dies from skin failure." Indeed, the skin faithfully renews itself throughout our lifetime and we usually take for granted its remarkable ability to expand during wound healing. This capacity for proliferation depends on stem cells located in hair follicles and in the epidermis, the outermost layer of the skin (see the figure) (1, 2). Stem cells in these areas are distinguished from surrounding epidermal cells by their essentially quiescent nature. Stem cells occasionally divide to maintain their presence in the skin (a process called "renewal") but they also generate rapidly proliferat-

ing cells called transient or transient amplifying cells. These cells are committed to differentiate and are necessary for tissue homeostasis and repair. Controlling the transition of a stem cell to a transient amplifying cell is of utmost importance to the skin. Too many transient amplifying cells may cause hyperplasia or thickening of the epidermis. On the other hand, skin thinning and atrophy can result from too few transient amplifying cells. These cells must themselves undergo terminal differentiation as they move to superficial positions in the skin before ultimately being sloughed from the surface. The well-orchestrated transitions between these cellular states require fine control, because alterations in these events could lead to cancer, nonmalignant growths, or poor tissue healing.

On page 933 of this issue, Aznar Benitah *et al.* (3) convincingly show that deletion of Rac1, a small GTP-binding protein of the Rho family, has a profound effect on skin homeostasis. After conditional deletion of Rac1 in mice, the

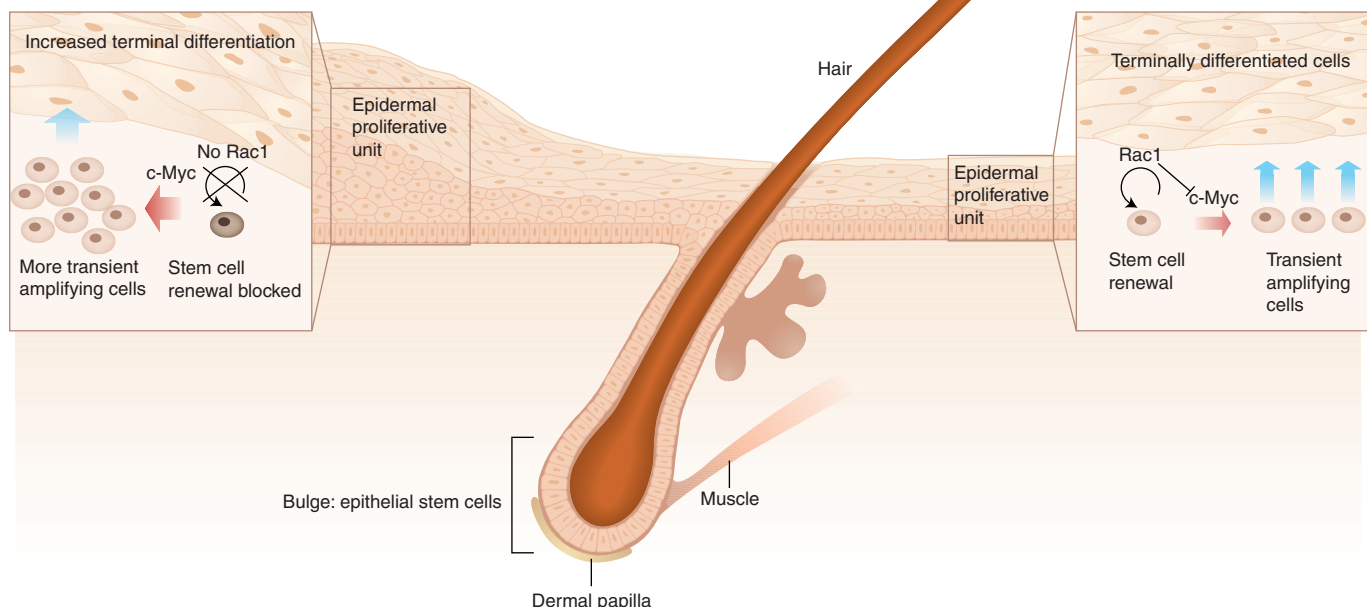
authors noted transient hyperproliferation and then premature terminal differentiation of epidermal cells. The epidermis first thickened, then thinned, and ultimately the skin failed to function properly. Mice developed cysts and hair loss. This series of events suggests that loss of Rac1 prevents stem cell renewal. This consequently would result in conversion of stem cells into transient amplifying cells, causing hyperplasia. The finding implies that in vivo, Rac1 normally maintains the epidermal stem cell niche by suppressing the switch that promotes the first transition event to committed cell populations.

A second complementary explanation for the skin-thinning phenotype is an increase in the later transition of transient amplifying cells to terminal differentiation. This could also result in depletion of the stem cell reservoir and is consistent with a dynamic equilibrium model of the stem cell, transient amplifying, and terminally differentiated cell populations (4).

The striking effect of Rac1 deletion on skin homeostasis is likely not mediated by effects on stem cells alone. Rac1 is expressed in all skin cells of the basal layer of the epidermis, not just the stem cells. One possibility is that expression of Rac1 is

G. P. Dotto is at the Institute of Biochemistry, University of Lausanne, Chemin de Bovares 155, CH-1066 Epalinges, Switzerland, and the Cutaneous Biology Research Center, Massachusetts General Hospital, Charlestown, MA 02129, USA. G. Cotsarelis is in the Department of Dermatology, Kligman Laboratories, University of Pennsylvania School of Medicine, M8 Stellar-Chance Building, 422 Curie Boulevard, Philadelphia, PA 19104, USA. E-mail: cotsarel@mail.med.upenn.edu

Skin thickening, followed by thinning, cysts, and hair loss, leading to skin failure



Epidermal stem cells require Rac1 protein. Stem cells in the mouse skin are localized to the epidermis and hair follicle. Within epidermal proliferative units, self-renewing stem cells give rise to transient amplifying cells. These then differentiate as they move upward in the skin. In normal skin, Rac1 promotes stem cell self-renewal. Rac1 inhibits stem cell differentiation, possibly by inhibiting c-Myc protein. Lack of Rac1 can cause an increase in committed cell populations, leading to skin failure.

vital for cell survival and that its loss simply caused a “toxic” effect in cells. This seems unlikely for two reasons. First, the orderly sequence of events after Rac1 deletion did not result in apoptosis or necrosis of epidermal cells. This indicates that loss of Rac1 has specific effects on cell proliferation and differentiation. Second, overexpression of Rac1 in cultured keratinocytes, the major constituent of the epidermis (constituting 90% of the tissue), resulted in increased cell proliferation. Overexpression studies in mice will be important to verify the proliferative and possibly tumorigenic effects of Rac1 *in vivo*.

These findings help us focus on what remains to be identified: the mechanisms upstream of Rac1 that control its function in the skin, and the downstream mediator(s) of Rac1 function in the control of stem cell renewal versus differentiation.

What is upstream of Rac1? The authors propose that niche conditions—more specifically, attachment of stem cells to the surrounding matrix—may signal to Rac1 and enhance its function in maintaining the quiescent stem cell population. A common mechanism for activation of small GTP-binding proteins involves GTP exchange factors (5), and an important candidate in this respect could be Tiam1. Mice lacking Tiam1 show decreased sensitivity to initiation of skin carcinogenesis (6), consistent

with the hypothesis that skin tumors originate from stem cells. Absence of Tiam1 would result in less Rac1 activation and therefore fewer stem cells.

What is downstream of Rac1? Rac1 regulates expression of the transcription factor c-Myc (7). Rac1 also controls c-Myc activity by regulating its phosphorylation by an enzyme called Pak (8). Aznar-Benitah *et al.* provide evidence supporting the latter mechanism in the skin. Previous findings indicated that increased c-Myc activity leads to transient epidermal hyperproliferation with long-term depletion of stem cell populations (9, 10). Thus, an elegantly linear model can be proposed whereby Rac1 activation keeps c-Myc activity in check, thereby suppressing stem cell depletion (see the figure). However, well-known cross-connections between Rac1 and other pathways with key roles in keratinocyte differentiation should be kept in mind.

Another small GTP-binding protein, RhoA, was recently found to suppress epidermal differentiation, and thus could exert equally important effects on the balance between self-renewing stem cells and committed epidermal cell populations (11). Similar to Rac1, RhoA expression is elevated in stem cell-rich areas of the skin (12). In fact, KyoT1/2, a target gene of RhoA signaling in keratinocytes (11), is also highly expressed in these regions (13). KyoT1/2

can inhibit epidermal stem cell differentiation by modulating processes including adhesion to the extracellular matrix and/or signaling by the Notch receptor.

Rac1 clearly has a critical role in maintenance of the mouse interfollicular epidermis. A similar role in human epidermis is likely, given its expression pattern and effects on cultured human keratinocytes. Thus, Rac1-controlled signaling pathways may be excellent targets for therapeutic intervention in processes such as cutaneous wounds, aging, and even hair loss.

References

1. R. J. Morris *et al.*, *Nat. Biotechnol.* **22**, 411 (2004).
2. C. S. Potten, C. Booth, *J. Invest. Dermatol.* **119**, 888 (2002).
3. S. Aznar Benitah, M. Frye, M. Glogauer, F. M. Watt, *Science* **309**, 933 (2005).
4. R. Okuyama, K. LeFort, G. P. Dotto, *J. Invest. Dermatol. Symp. Proc.* **9**, 248 (2004).
5. K. Burridge, K. Wennerberg, *Cell* **116**, 167 (2004).
6. A. Malliri *et al.*, *Nature* **417**, 867 (2002).
7. M. Chiariello, M. J. Marinissen, J. S. Gutkind, *Nat. Cell Biol.* **3**, 580 (2001).
8. Z. Huang, J. A. Traugh, J. M. Bishop, *Mol. Cell. Biol.* **24**, 1582 (2004).
9. I. Arnold, F. M. Watt, *Curr. Biol.* **11**, 558 (2001).
10. R. L. Waikel, Y. Kawachi, P. A. Waikel, X. J. Wang, D. R. Roop, *Nat. Genet.* **28**, 165 (2001).
11. M. Grossi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
12. X. Xu, S. Lyle, Y. Liu, B. Solky, G. Cotsarelis, *Am. J. Pathol.* **163**, 969 (2003).
13. T. Tumber *et al.*, *Science* **303**, 359 (2004).

10.1126/science.1117192

The Coming Paradigm Shift in Forensic Identification Science

Michael J. Saks¹ and Jonathan J. Koehler²

Converging legal and scientific forces are pushing the traditional forensic identification sciences toward fundamental change. The assumption of discernible uniqueness that resides at the core of these fields is weakened by evidence of errors in proficiency testing and in actual cases. Changes in the law pertaining to the admissibility of expert evidence in court, together with the emergence of DNA typing as a model for a scientifically defensible approach to questions of shared identity, are driving the older forensic sciences toward a new scientific paradigm.

Little more than a decade ago, forensic individualization scientists compared pairs of marks (handwriting, fingerprints, tool marks, hair, tire marks, bite marks, etc.), intuited whether the marks matched, and testified in court that whoever or whatever made one made the other. Courts almost never excluded the testimony. Cross-examination rarely questioned the foundations of the asserted expertise or the basis of the analyst's certainty.

Today, that once-complacent corner of the law and science interface has begun to unravel—or at least to regroup. The news carries reports of erroneous forensic identifications of hair, bullets, handwriting, footprints, bite marks, and even venerated fingerprints. Scientists have begun to question the core assumptions of numerous forensic sciences (1–6). Federal funding has materialized to support research that examines long-asserted but unproven claims. Courts have started taking challenges to asserted forensic science expertise seriously (1). A dispassionate scientist or judge reviewing the current state of the traditional forensic sciences would likely regard their claims as plausible, underresearched, and oversold.

The traditional forensic individualization sciences rest on a central assumption: that two indistinguishable marks must have been produced by a single object. Traditional forensic scientists seek to link crime scene evidence to a single person or object “to the exclusion of all others in the world” (7, 8). They do so by leaning on the assumption of discernible uniqueness. According to this assumption, markings produced by different people or objects are observably different. Thus, when a pair of markings is not observably

different, criminalists conclude that the marks were made by the same person or object.

Although lacking theoretical or empirical foundations, the assumption of discernible uniqueness offers important practical benefits to the traditional forensic sciences. It enables forensic scientists to draw bold, definitive conclusions that can make or break cases. It excuses the forensic sciences from developing measures of object attributes, collecting population data on the frequencies of variations in those attributes, testing attribute independence, or calculating and explaining the probability

that different objects share a common set of observable attributes. Without the discernible uniqueness assumption, far more scientific work would be needed, and criminalists would need to offer more tempered opinions in court.

Legal and scientific forces are converging to drive an emerging skepticism about the claims of the traditional forensic individualization sciences. As a result, these sciences are moving toward a new scientific paradigm. [We use the notion of paradigm shift not as a literal application of Thomas Kuhn's concept (9), but as a metaphor highlighting the transformation involved in moving from a pre-science to an empirically grounded science.] Two such forces are outgrowths of DNA typing: the discovery of erroneous convictions and a model for a scientifically sound identification science. A third force is the momentous change in the legal admissibility standards for expert testimony. A final force grows from studies of error rates across the forensic sciences.

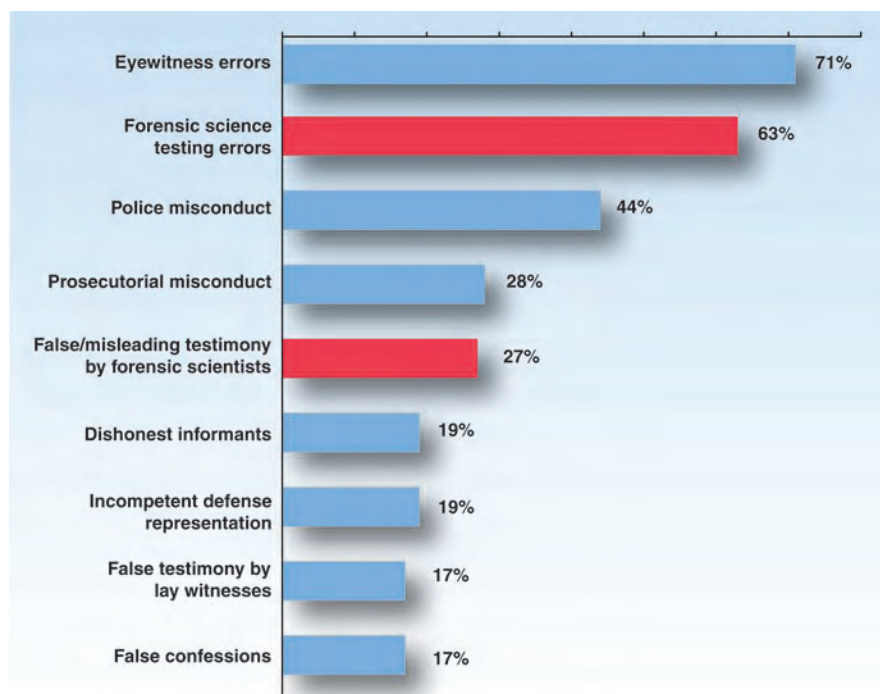


Fig. 1. Factors associated with wrongful conviction in 86 DNA exoneration cases, based on case analysis data provided by the Innocence Project, Cardozo School of Law (New York, NY), and computed by us. Percentages exceed 100% because more than one factor was found in many cases. Red bars indicate factors related to forensic science.

¹College of Law, Arizona State University, Tempe, AZ 85287, USA. E-mail: saks@asu.edu ²McCombs School of Business, University of Texas, Austin, TX 78712, USA. E-mail: koehler@mail.utexas.edu



Fig. 2. Bite mark evidence exhibit from trial of Ray Krone, suggesting alignment of a cast of Krone's dentition with bite wounds in victim's flesh [*State v. Krone*, 182 Ariz. 319 (1995)]. A forensic odontologist testified that this showed Krone to be the biter. Krone was convicted of murder and sentenced to death, but a decade later he was exonerated by DNA analysis. [Source: E. Thomas Barham (Los Alamitos, CA) and Alan Simpson (Phoenix, AZ), attorneys for Krone]

Post-Conviction DNA Exonerations

During the past decade, scores of people who were convicted of serious crimes—including at least 14 who had been sentenced to death—have been exonerated by DNA analyses of crime scene evidence that had not been tested at the time of their trials (10). It was not surprising to learn that erroneous convictions sometimes occur, and that new science and technology can help detect and correct those mistakes. Nor was it surprising to learn, from an analysis of 86 such cases (Fig. 1), that erroneous eyewitness identifications are the most common contributing factor to wrongful convictions. What was unexpected is that erroneous forensic science expert testimony is the second most common contributing factor to wrongful convictions, found in 63% of those cases. These data likely understate the relative contribution of forensic science expert testimony to erroneous convictions. Whereas lawyers, police, and lay witnesses participate in virtually every criminal case, forensic science experts participate in a smaller subset of cases—about 10 to 20% of criminal cases during the era when these DNA exonerations were originally tried (11).

Figure 1 also indicates that forensic scientists are the witnesses most likely to present misleading or fraudulent testimony. Deceitful forensic scientists are a minor sidelight to this paper, but a sidelight that underscores cultural differences between normal science and forensic science (12, 13). In normal science, academically gifted students receive four or more years of doctoral training where much of the socialization into the culture of science takes place. This culture emphasizes methodological rigor, openness, and cautious interpretation of data. In forensic science, 96% of positions are held by persons with bachelor's degrees (or less), 3% master's degrees, and 1% Ph.D.s (14). When individuals who are not steeped in the culture of science work in an adversarial, crime-fighting culture, there is a substantial risk that a different set of norms will prevail. As one

former forensic scientist noted, this pressure-packed environment can lead to data fudging and fabrication: "All [forensic science] experts are tempted, many times in their careers, to report positive results when their inquiries come up inconclusive, or indeed to report a negative result as positive" [(15), p. 17].

DNA Typing as the New Model for Scientific Forensic Identification

Much of the above criticism does not apply to the science of DNA typing as practiced today. Indeed, DNA typing can serve as a model for the traditional forensic sciences in three important respects. First, DNA typing technology was an application of knowledge derived from core scientific disciplines. This provided a stable structure for future empirical work on the technology. Second, the courts and scientists scrutinized applications of the technology in individual cases. As a result, early, unscientific practices were rooted out. Third, DNA typing offered data-based, probabilistic assessments of the meaning of evidentiary "matches." This practice represented an advance over potentially misleading match/no-match claims associated with other forensic identification sciences.

Immediately after DNA's first courtroom appearance in the 1980s, scientists from disciplines as varied as statistics, psychology, and evolutionary biology debated the strengths and limitations of forensic DNA evidence. Blue-ribbon panels were convened, conferences were held, unscientific practices were identified, data were collected, critical papers were written, and standards were developed and implemented. The scientific debates focused on the adequacy of DNA databases (16), the computation of DNA match probabilities (17), the training of DNA analysts (18), the presentation of DNA matches in the courtroom (19), and the role of error rates (20). In some cases, disputants worked together to find common ground (21). These matters were not resolved by the forensic scientists themselves, by fiat, or by neglect. Most exaggerated claims and counterclaims about DNA evidence have been replaced by scientifically defensible propositions. Although some disagreement remains (22), the scientific process worked.

One of the great strengths of DNA typing is that it uses a statistical approach based on population genetics theory and empirical testing. Experts evaluate matches between suspects and crime scene DNA evidence in terms of the probability of random matches across different reference populations (e.g., different ethnicities). These probabilities are derived from databases that identify the frequency with which var-

ious alleles occur at different locations on the DNA strand. The traditional forensic sciences could and should emulate this approach (23). Each subfield must construct databases of sample characteristics and use these databases to support a probabilistic approach to identification. Fingerprinting could be one of the first areas to make the transition to this approach because large fingerprint databases already exist. The greatest challenge in this effort would be to develop measures of the complex images presented by fingerprints, tool marks, bite marks, handwriting, etc. (Figs. 2 and 3). Forensic scientists will need to work with experts in differential geometry, topology, or other fields to develop workable measures.

A second data collection effort that would strengthen the scientific foundation of the forensic sciences involves estimating error rates. Although the theoretical promise of forensic technology is considerable, the practical value of any particular technology is limited by the extent to which potentially important errors arise. The best way to identify the frequency with which errors occur is to conduct blind, external proficiency tests using realistic samples. A proficiency test requires analysts to make judgments about samples whose properties are known. External proficiency tests are conducted by an agency unaffiliated with the forensic scientist's laboratory. Externality is important to the integrity of proficiency tests because laboratories have strong incentives to

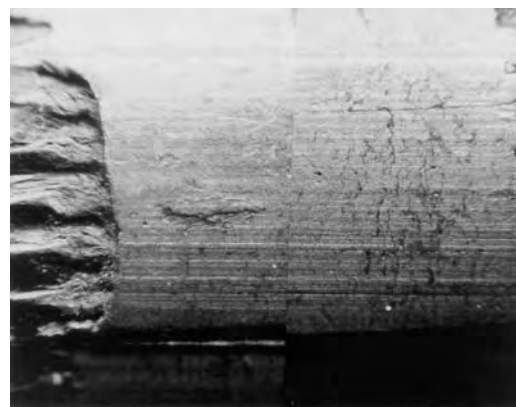


Fig. 3. Image of two bullets viewed through a comparison microscope. The bullets were fired from two consecutively manufactured Smith & Wesson 38 Special revolver barrels. Whether fired through the same or different barrels, numerous matching and nonmatching striations are engraved onto bullets. To reliably identify the barrel through which a questioned bullet was fired, an examiner must distinguish among class, subclass, and individual characteristics. These two bullets illustrate subclass characteristic agreement of striated markings on a groove impression that could be mistaken for individual characteristics. Without investigating the potential for subclass carryover, the examiner could mistake these as having been fired from the same gun. [Source: Bruce Moran, firearms examiner with the Sacramento County (CA) District Attorney, Laboratory of Forensic Services]

be perceived as error-free. An even better test would be a blind proficiency test, in which the analyst believes the test materials are part of ordinary case work. Blindness increases the validity of proficiency test results because it ensures that analysts treat the test sample as they would other case samples. Although proficiency tests are used in many forensic sciences, the tests are generally infrequent, internal, and unrealistic; blind tests are practically nonexistent.

Changes in the Law

Until recently, courts assessed expertise by looking for superficial indicia of validity. In the 19th century, courts were impressed by “qualifications” and success in the marketplace. If the market valued an asserted expertise or expert, courts generally did, too. In *Frye v. United States* [293 F. 1013 (D.C. Cir. 1923)], a federal appellate court confronted the question of admissibility of an expertise that had no life in any commercial marketplace. The court solved the problem by substituting an intellectual marketplace. The court asked whether the proffered expertise had “gained general acceptance in the particular field in which it belongs.” Sixty years later, the *Frye* test had become the dominant expert evidence filter in American courts.

In 1993, the law began to catch up with the scientific method. In *Daubert v. Merrell Dow Pharmaceuticals* [509 U.S. 579 (1993)], the U.S. Supreme Court introduced a new standard for the admissibility of scientific evidence. Under *Daubert*, proffered scientific testimony must be shown to stand on a dependable foundation. The court suggested that trial judges making this determination consider whether the proffered science has been tested, the methodological soundness of that testing, and the results of that testing. The *Daubert* test in effect lowers the threshold for admission of sound cutting-edge science and raises the threshold for long-asserted expertise that lacks a scientific foundation. Seriously applied, the *Daubert* test subjects the forensic sciences to a first-principles scientific scrutiny that poses a profound challenge to fields that lack rigorous supporting data.

United States v. Starzecpyzel [880 F. Supp. 1027 (S.D.N.Y. 1995)] offered an early indication of how *Daubert* could change judicial views. After an extensive hearing on the soundness of asserted handwriting identification expertise, a federal district court concluded that the field had no scientific basis: “[T]he testimony at the *Daubert* hearing firmly established that forensic document examination, despite the existence of a certification program, professional journals and other trappings

of science, cannot, after *Daubert*, be regarded as ‘scientific ... knowledge’” (p. 1038). However, the court did not exclude this unscientific testimony. It reasoned that handwriting identification did not have to reach the *Daubert* standard because *Daubert* applied only to scientific evidence, and handwriting identification plainly was not scientific evidence. Thus, when a forensic science was found to stand on a weak foundation, the threshold of admission was lowered to accommodate this weakness.

In *Kumho Tire v. Carmichael* [526 U.S. 137 (1999)], the Supreme Court directly confronted the question of whether *Daubert* applies to nonsciences. A consortium of law enforcement organizations prepared an amicus brief urging that *Daubert* scrutiny not be extended to the testimony of police agency expert witnesses. The brief argued that “the great bulk of expert testimony provided by law enforcement officers does not involve sci-

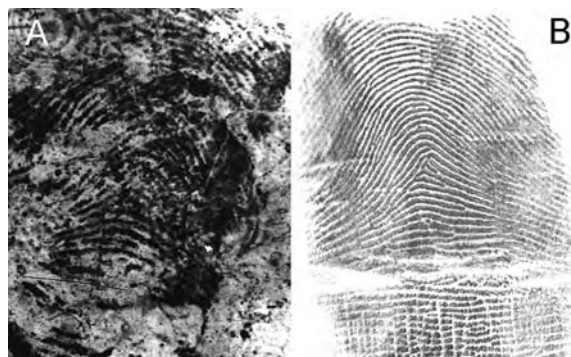


Fig. 4. (A) A latent fingerprint believed to belong to a terrorist involved in train bombings in Madrid, Spain, in March 2004. (B) A database print belonging to Brandon Mayfield of Portland, Oregon. On the basis of these prints (though not necessarily these very images), FBI fingerprint examiners erroneously identified Mayfield as the bomber (26). [Source: Problem Idents, onin.com/fp/problemidents.html#madrid]

entific theories, methodologies, techniques, or data in any respect.... Instead, law enforcement officers testify about such things as accident reconstruction, fingerprint, footprint and handprint [identification], handwriting analysis, firearms markings and toolmarks and the unique characteristics of guns, bullets, and shell casings, and bloodstain pattern identification” (24). Ironically, then, fields that initially gained entry to the courts by declaring themselves to be “sciences” now sought to remain in court by denying any connection with scientific methods, data, or principles. Despite efforts to preserve the “nonscience” loophole, the Supreme Court doctrinally sealed it shut when *Kumho Tire* held that all expert testimony must pass appropriate tests of validity to be admissible in court.

Error Rates

Although *Daubert*’s testing recommendations are familiar to most scientists, there has

been remarkably little research on the accuracy of traditional forensic sciences. Proficiency tests in some fields offer a step in the right direction, even though simple tasks and infrequent peer review limit their value. Nonetheless, the available data hint that some forensic sciences are best interpreted in tandem with error rates estimated from sound studies.

Unfortunately, forensic scientists often reject error rate estimates in favor of arguments that theirs is an error-free science. For example, an FBI document section chief asserted that all certified document examiners in the United States would agree with his conclusions in every case [(25), p. 196]. Likewise, fingerprint experts commonly claim that all fingerprint experts would reach the same conclusions about every print (2). Such hubris was on display in spring 2004 when the FBI declared that a fingerprint recovered from a suspicious plastic bag near the scene of a terrorist bombing in Madrid provided a “100 percent match” to an Oregon attorney (Fig. 4). The FBI eventually conceded error when Spanish fingerprint experts linked the print to someone else (26).

The FBI and other agencies often seek to preserve the illusion of perfection after disclosure of such errors by distinguishing between human errors (“possible”) and errors of method (“impossible”). A leading FBI scientist explained the distinction to the court in *United States v. Llera-Plaza I* [58 Fed. R. Evid. Serv. 1 (E.D. Pa. 2002)]: “We have to understand that error rate is a difficult thing to calculate. I mean, people are trying to do this, it shouldn’t be done, it can’t be done.... An error rate is a wispy thing like smoke, it changes over time.... If you made a mistake in the past, certainly that’s valid information ... but to say there’s an error rate that’s definable would be a misrepresentation.... Now, error rate deals with people, you should have a method that is defined and stays within its limits, so it doesn’t have error at all. So the method is one thing, people making mistakes is another issue.”

Such claims are problematic. First, the suggestion that humans err but forensic techniques do not is unfalsifiable. It is impossible to disentangle “method” errors from “practitioner” errors in fields where the method is primarily the judgment of the examiner. Second, even if such disentanglement were possible, it is a red herring. When fact-finders hear evidence of a forensic match, a proper assessment of the probative value of that match requires awareness of the chance that a mistake was made. The source of such a mistake is irrelevant for this purpose. If method errors could be distinguished from practitioner errors,

a 1% method error affects the probative value of the match in exactly the same way as a 1% practitioner error. Identifying sources of error is relevant for improving forensic science practice, but it plays no role in identifying the probative importance of a match.

Third, the suggestion that error rates do not exist because they change over time and are not specific to the case at hand is a base-rate fallacy. In this fallacy of reasoning, people underuse (or willfully ignore) general background data in judgment tasks because they believe the data are irrelevant to the instant case. However, general background data (or base rates) are relevant for specific predictions (27, 28). For example, although risk estimates for a disease fluctuate and are developed on patients other than the patient now seeking medical advice, these estimates provide information useful for predicting whether this patient will contract the disease. A 20% base-rate risk of contracting the disease makes it more likely that the patient will get the disease than would a 1% risk. Likewise, an X% base-rate risk of error in a given forensic science provides some indication of the chance that a particular conclusion is in error (22).

Data from proficiency tests and other examinations suggest that forensic errors are not minor imperfections. Spectrographic voice identification error rates are as high as 63%, depending on the type of voice sample tested [(1), chap. 31]. Handwriting error rates average around 40% and sometimes approach 100% [(1), chap. 28]. False-positive error rates for bite marks run as high as 64% [(1), chap. 30]. Those for microscopic hair comparisons are about 12% (using results of mitochondrial DNA testing as the criterion) (29). Fingerprint examiners generally fare better, although data from a well-known forensic testing program contradict industry boasts of perfect, or even near-perfect, agreement (30). Since 1995, about one-fourth of examiners failed to correctly identify all latent prints in this test (which includes 9 to 12 latent prints and palmprints). About 4 to 5% of examiners committed false-positive errors on at least one latent. In one test, 20% of examiners mistook one person's prints for those of his twin. The editor of the leading fingerprint journal called this performance "unacceptable" [(31), p. 524]. It is noteworthy that these misidentifications are not confined to a single lab, circumstance, or marking. Moreover, the misidentification rates do not show a clear pattern of improvement (the misidentification rates in 2004 were 4 to 6%). Nor are these errors limited to arguably arti-

ficial testing situations; erroneous fingerprint identifications have made their way out of the crime lab and into prosecutions in at least 21 documented cases (32).

Forensic science proficiency tests and examinations are obviously imperfect indicators of the rate at which errors occur in practice. This fact does not justify ignoring the worrisome data these tests have yielded. Indeed, these data are probably best regarded as lower-bound estimates of error rates. Because the tests are relatively easy (according to test participants), and because participants know that mistakes will be identified and punished, test error rates (particularly the false-positive error rate) probably are lower than those in everyday casework (33, 34).

The studies mentioned above cry out for attention and follow-up investigations. In light of the law's growing reluctance to accept experts' personal guarantees in lieu of scientific data, these studies should increase candor about performance and create pressure for improvement.

The Future

The traditional forensic sciences need look no further than their newest sister discipline, DNA typing, for guidance on how to put the science into forensic identification science. This effort should begin with adoption of the basic-research model. Just as DNA scientists tested the genetic assumptions that undergirded DNA typing theory (e.g., Hardy-Weinberg equilibrium), traditional forensic scientists should design experiments that test the core assumptions of their fields. As basic research knowledge grows, experts will be able to inform courts about the relative strengths and weaknesses of their theories and methods, and suggest how that knowledge applies to individual cases.

At the same time, data should be collected on the frequency with which markings and attribute variations occur in different populations. In addition to their case-specific benefits, these data may also facilitate the development of artificial intelligence or computer-aided pattern recognition programs for the identification sciences. Forensic scientists might also adopt protocols, such as blind examinations in combination with realistic samples, that minimize the risks that their success rates will be inflated and their conclusions biased by extraneous evidence and assumptions (34). When matches are identified, forensic scientists in all fields would compute and report random-match probabilities similar to those used in DNA typing. These estimates—in combination with error rate estimates provided by mandatory, well-

constructed proficiency tests—would inform fact-finders about the probative value of the evidentiary match.

Simply put, we envision a paradigm shift in the traditional forensic identification sciences in which untested assumptions and semi-informed guesswork are replaced by a sound scientific foundation and justifiable protocols. Although obstacles exist both inside and outside forensic science, the time is ripe for the traditional forensic sciences to replace antiquated assumptions of uniqueness and perfection with a more defensible empirical and probabilistic foundation.

References

1. D. L. Faigman, D. Kaye, M. J. Saks, J. Sanders, *Modern Scientific Evidence: The Law and Science of Expert Testimony* (Thompson-West, St. Paul, MN, ed. 2, 2002).
2. R. Epstein, *S. Cal. Law Rev.* **75**, 605 (2002).
3. M. J. Saks, *Hastings Law J.* **49**, 1069 (1998).
4. National Research Council, *Forensic Analysis: Weighing Bullet Lead Evidence* (National Academy Press, Washington, DC, 2004).
5. I. Pretty, D. Sweet, *Sci. Justice* **41**, 85 (2001).
6. C. M. Cooley, *Stanford Law Policy Rev.* **15**, 381 (2004).
7. D. A. Stoney, *J. Forensic Sci. Soc.* **31**, 197 (1991).
8. J. W. Osterburg, *J. Crim. Law Criminol.* **60**, 97 (1969).
9. T. S. Kuhn, *The Structure of Scientific Revolutions* (Univ. of Chicago Press, Chicago, 1962).
10. B. Scheck, P. Neufeld, J. Dwyer, *Actual Innocence* (Doubleday, New York, 2000).
11. J. Peterson, *Forensic Science and the Courts* (National Institute of Justice, Washington, DC, January 1986).
12. P. C. Giannelli, *Issues Sci. Tech.* **20**, 55 (fall 2003).
13. P. C. Giannelli, *Va. J. Soc. Policy Law* **4**, 439 (1997).
14. K. G. Furtton, Y. L. Hsu, M. D. Cole, *J. Forensic Sci.* **44**, 128 (1999).
15. A. A. Moenssens, *J. Crim. Law Criminol.* **84**, 1 (1993).
16. Committee on DNA Technology in Forensic Science, *DNA Technology in Forensic Science* (National Research Council, Washington, DC, 1992).
17. R. C. Lewontin, D. L. Hartl, *Science* **254**, 1745 (1991).
18. Technical Working Group on DNA Analysis Methods, *Crime Lab. Digest* **18**, 44 (1991).
19. K. Roeder, *Stat. Sci.* **9**, 263 (1994).
20. J. J. Koehler, A. Chia, J. S. Lindsey, *Jurimetrics J.* **35**, 201 (1995).
21. B. Budowle, E. S. Lander, *Nature* **371**, 735 (1994).
22. J. J. Koehler, *Jurimetrics J.* **37**, 425 (1997).
23. M. J. Saks, J. J. Koehler, *Cardozo Law Rev.* **13**, 361 (1991).
24. Brief Amicus Curiae of Americans for Effective Law Enforcement et al., submitted in *Kumho Tire v. Carmichael*, No. 97-1709, in the Supreme Court of the United States (October Term, 1997).
25. D. Fisher, *Hard Evidence* (Simon & Schuster, New York, 1995).
26. D. Heath, H. Bernton, Portland lawyer released in probe of Spain bombings, *Seattle Times* (21 May 2004).
27. D. Kahneman, A. Tversky, *Psychol. Rev.* **80**, 237 (1973).
28. J. J. Koehler, *Behav. Brain Sci.* **19**, 1 (1996).
29. M. M. Houck, B. Budowle, *J. Forensic Sci.* **47**, 964 (2002).
30. Collaborative Testing Services Inc., *Forensics Testing Program: Latent Prints Examination, 1993-2004*.
31. D. Grieve, *J. Forensic Ident.* **46**, 524 (1996).
32. S. Cole, *J. Crim. Law Criminol.*, in press (available at www.seweb.uci.edu/faculty/cole).
33. J. L. Peterson et al., *Forensic Evidence and the Police* (National Institute of Justice, Washington, DC, 1984).
34. M. J. Saks, D. M. Risinger, R. Rosenthal, W. Thompson, *Sci. Justice* **43**, 77 (2003).

10.1126/science.1111565

Extreme Waves Under Hurricane Ivan

David W. Wang,* Douglas A. Mitchell, William J. Teague, Ewa Jarosz, Mark S. Hulbert

On 15 September 2004, the center of Hurricane Ivan (Fig. 1A and fig. S1) passed directly over six wave-tide gauges deployed by the Naval Research Laboratory (NRL), at depths of 60 and 90 m, on the outer continental shelf in the northeastern Gulf of Mexico, allowing us to measure the extreme waves directly under a category 4 hurricane (1). We calculated significant wave height (H_s) and maximum individual wave height (H_{\max}), two parameters commonly used to characterize wave fields (2).

During Ivan's approach, H_s and H_{\max} rapidly increased and reached peak values when the radial distance between the eye's center and the moorings was ~ 75 km (Fig. 1B). H_s reached maximum values of 17.9, 16.1, and 17.1 m at moorings 3, 4, and 5, respectively. These H_s values were larger than those measured the same day by National Data Buoy Center (NDBC) buoy 42040 (Fig. 1A), which recorded the largest H_s (15.96 m) ever reported by NDBC. The largest H_{\max} reached 27.7 m (91 ft) at mooring 3; out of 146 waves measured at

moorings 3, 4, and 5, there were 24 individual waves with heights greater than 15 m (50 ft) (1).

The measured values of H_s and H_{\max} depict the radial variability of the hurricane wave field in the range $1 \leq r/R \leq 8$ (Fig. 1C), where r is the radial distance from the moorings to the eye's center and R is the radius of maximum winds (40 km) (3). H_s increased rapidly as the normalized radial distance approached 1 (Fig. 1, B and C) and can be approximated by an exponential curve of the form $H_s = a(r/R)^b \exp[-(r/R)^c]$, where $a = 56.61$ m, $b = -0.96$, and $c = -0.94$ (Eq. 1). This compares well with a numerical model (4), provided the model's H_s is set to 21 m at $r/R = 1$ (Fig. 1C). Past observations of H_{\max} during hurricane-generated seas suggest that H_{\max} can reach $1.9H_s$ (5), which is consistent with the upper limit of our measurements (Fig. 1B).

The wave-sampling strategy (1) employed captured a small segment of the wave field, suggesting our measurements likely missed the largest waves near the storm's eyewall. The largest

measured H_s reached 17.9 m at a radial distance of 73 km, about 30 km from the strongest winds. Furthermore, our measurements, from the forward face of Ivan, are likely $\sim 85\%$ of the maximum H_s typically found in the right quadrant (4, 6). These factors strongly suggest the wave field associated with Ivan should generate maximum H_s values greater than 21 m and H_{\max} values greater than 40 m at $r/R = 1$.

The values of H_s measured here, possibly reduced by shoaling, are larger than those predicted by several parametric wave models developed for deep water conditions. Young (6) proposed a semi-empirical model based on R , maximum wind speed (U_{\max}), and hurricane translation speed (V_t); with $R = 40$ km, $V_t = 6$ m s $^{-1}$, and $U_{\max} = 60$ m s $^{-1}$, the model predicts a maximum H_s of 15.1 m. Hsu (7) suggested a simple empirically determined formula, $H_s = 0.2(P_R - P_0)$, where $P_R = 1013$ mbar is the pressure at the edge of the hurricane and $P_0 = 935$ mbar is the central pressure, resulting in an H_s of 15.6 m. Underestimation by these models likely stems from the absence of wave data under intense storms. Measurements of the extremely large waves directly under Ivan may act as a starting point for improving our understanding of the waves generated by the most powerful hurricanes.

References and Notes

1. Materials and methods are available as supporting material on Science Online.
2. M. J. Tucker, *Waves in Ocean Engineering: Measurement, Analysis, Interpretations* (Ellis Horwood, Chichester, UK, 1991).
3. Hurricane Research Division, Atlantic Oceanographic and Meteorological Laboratory, "Hurricane Ivan 2004" (available at www.aoml.noaa.gov/hrd/Storm_pages/ivan2004/).
4. Coastal Engineering Research Center, U.S. Army Corps of Engineers, *Shore Protection Manual* (Government Printing Office, Washington, DC, 1984), vol. 1.
5. R. G. Bea, in *Proceedings of the Sixth Offshore Technology Conference*, Houston, TX, 76 to 8 May 1974 (Offshore Technology Conference, Dallas, TX, 1974), pp. 791–810.
6. I. R. Young, *J. Wtrwy. Port Coast. Ocean Eng. ASCE* **114**, 637 (1988).
7. S. A. Hsu, M. F. Martin Jr., B. W. Blanchard, *J. Coastal Res.* **16**, 832 (2000).
8. Supported by the Office of Naval Research as part of the NRL's basic research project "Slope to Shelf Energetics and Exchange Dynamics (SEED)" under program element 0601153N (NRL–Stennis Space Center contribution no. NRL/JA/7330–05–5172), by the Minerals Management Service Environmental Studies Program, and by the Minerals Management Service Technology Assessment and Research Program on Hurricane Ivan.

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5736/896/DC1

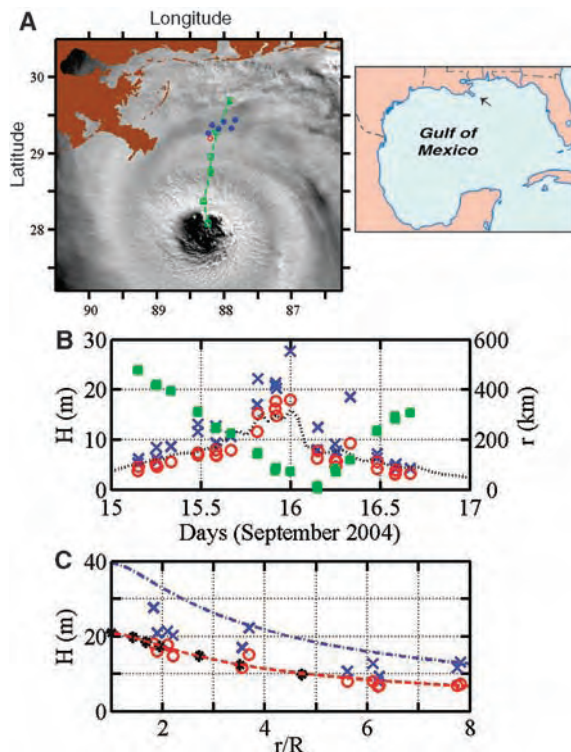
Materials and Methods
Fig. S1

21 March 2005; accepted 8 July 2005
10.1126/science.1112509

Naval Research Laboratory, Stennis Space Center, MS 39529, USA.

*To whom correspondence should be addressed.
E-mail: dwang@nrlssc.navy.mil

Fig. 1. (A) Satellite image of Hurricane Ivan from the Moderate Resolution Imaging Spectroradiometer (MODIS) at 1850 universal time, 15 September 2004 (provided by NRL's Ocean Optics Group). The eye of Hurricane Ivan is clearly shown just southeast of the boot of Louisiana. NRL moorings are shown as blue dots [northern line (60 m), moorings 1, 2, and 3; southern line (90 m), moorings 4, 5, and 6]. The NDBC buoy is shown as a red circle, and the track of Hurricane Ivan is shown as a green dashed line with squares marking the hurricane's center every 3 hours. (Inset) Location of Ivan at the time of measurement. (B) Time evolution of H_s (circles) and H_{\max} (crosses) for the six NRL moorings, H_s for NDBC buoy 42040 (dotted line), and radial distance to Ivan's center (squares). (C) H_s and H_{\max} as a function of normalized radial distance (r/R). The red dashed line represents the exponential relation (Eq. 1); digitized values of a segment 15° clockwise from the forward direction of a numerically simulated wave field are denoted by black asterisks. The blue dashed line represents $H_{\max} = 1.9H_s$, and circles and crosses are as in (B).



Crystal Structure of a Mammalian Voltage-Dependent *Shaker* Family K⁺ Channel

Stephen B. Long, Ernest B. Campbell, Roderick MacKinnon*

Voltage-dependent potassium ion (K⁺) channels (Kv channels) conduct K⁺ ions across the cell membrane in response to changes in the membrane voltage, thereby regulating neuronal excitability by modulating the shape and frequency of action potentials. Here we report the crystal structure, at a resolution of 2.9 angstroms, of a mammalian Kv channel, Kv1.2, which is a member of the *Shaker* K⁺ channel family. This structure is in complex with an oxido-reductase β subunit of the kind that can regulate mammalian Kv channels in their native cell environment. The activation gate of the pore is open. Large side portals communicate between the pore and the cytoplasm. Electrostatic properties of the side portals and positions of the T1 domain and β subunit are consistent with electrophysiological studies of inactivation gating and with the possibility of K⁺ channel regulation by the β subunit.

Voltage-dependent K⁺ (Kv) channels are members of the voltage-dependent cation channel family, which includes the voltage-dependent K⁺, Na⁺, and Ca²⁺ channels (1). These channels are present in all the major kingdoms of life. In eukaryotic cells, they work in concert with other ion channels to produce and modulate the electrical activity of the cell. This electrical activity is important for many processes in electrically excitable cells such as neurons and muscle, as well as in nonexcitable cells (1). In the quintessential excitable cell, the neuron, Kv channels return the membrane voltage to its negative resting value after an action potential, modulate the shape of action potentials, and set the action potential firing rate (1).

Most of our knowledge of Kv channel function comes from studies of the *Shaker* K⁺ channel from *Drosophila melanogaster* and its family members from mammalian cells (2). *Shaker* family channels have been extensively studied with electrophysiology, because they can easily be expressed in *Xenopus laevis* oocytes and in other cells. In contrast, nearly all of our knowledge of K⁺ channel structure is based on studies of prokaryotic K⁺ channels, because they are more easily expressed at high levels in *Escherichia coli*. Such studies have taught us much about their pores, selectivity filters, and gates (3).

Eukaryotic Kv channels in many respects are very similar to their prokaryotic counter-

parts. The selectivity filter sequence is so conserved that we expect its structure to be essentially the same in all K⁺ channels. The pore's "inverted teepee" arrangement of inner helices, which holds the selectivity filter in its wider half near the extracellular surface, is also expected to be a conserved feature (4). However, beyond their conserved pore and certain domains that regulate the opening of the pore's gate, eukaryotic Kv channels have certain unique features. For example, in the S6 inner helix (on the intracellular side of the selectivity filter), a highly conserved triplet sequence, Pro-X-Pro (where X is any amino acid), is present in *Shaker* family Kv channels but not in prokaryotic Kv channels. Mutations show that this sequence is very important for gating, but the reason why has yet to be determined (5, 6). There has been speculation that this region of the pore (the inner pore), which is lined by the S6 inner helices, is different in *Shaker* Kv channels (5, 7) than in prokaryotes.

Shaker family Kv channels have an adaptation that, as far as we know, does not exist in prokaryotic Kv channels, and apparently allows them to carry out tasks that are unique to eukaryotic cells. Preceding the first membrane-spanning helix, S1, the N terminus forms a T1 domain inside the cell (8–11). Four T1 domains, one from each of the four *Shaker* Kv channel subunits, come together to form a tetrameric assembly at the intracellular membrane surface. This domain is located directly over the pore entryway to the cytoplasm, which means that the transmembrane pore must communicate with the cytoplasm through side portals in order to allow K⁺ ions to flow freely between the cell and the transmembrane pore (12–14). These portals not only permit

K⁺ ions, but they also must be large enough to allow the entry of a polypeptide chain from the channel's N terminus, which functions as an inactivation gate in some *Shaker* family Kv channels (15, 16).

The T1 domain in eukaryotic Kv channels is a docking platform for the β subunit (12, 17, 18). The β subunit forms a tetramer of proteins related to aldo-keto reductase enzymes, which are oxido-reductases dependent on NADPH (the reduced form of nicotinamide adenine dinucleotide phosphate), with α - β barrel structures (19–24). Crystal structures of a β subunit tetramer and of a β -T1 domain complex showed that the enzyme's active site contains an NADP⁺ cofactor and catalytic residues for hydride transfer (12, 24). The β subunit active site is mysterious because its function is still unknown. Is it an enzyme that can be regulated by a Kv channel or does it serve as a sensor for the Kv channel, allowing the redox state of a cell to influence electrical activity at the membrane? Nearly all *Shaker* family Kv channels in the mammalian nervous system are associated with these oxido-reductase β subunits (25).

Experimental evidence suggests that voltage-dependent gating is fundamentally similar in prokaryotic and eukaryotic Kv channels and that their voltage sensors are structurally similar (26–29). But structural studies with *Shaker* family Kv channels have the potential to be more informative. The structures of the prokaryotic Kv channel KvAP suggest that an intact lipid membrane is required to keep the voltage sensors correctly oriented with respect to the pore (30) [see also Protein Data Bank (PDB) ID 2A0L]. Because the membrane is removed during isolation of the channel, this requirement has so far obscured all attempts to deduce the structural basis for electro-mechanical coupling between the voltage sensor and the pore. The voltage sensor in *Shaker* family Kv channels is constrained by its attachment to a cytoplasmic T1 domain, and therefore may maintain its tenuous but important connections to the pore.

Structure of the Kv1.2- β subunit complex. The Kv1.2 K⁺ channel from rat brain (31) was coexpressed with the β 2 K⁺ channel β subunit from rat brain (32) in the yeast *Pichia pastoris* as described in (33) with modifications (34). The protein was purified and crystallized (34). Throughout purification and crystallization, it was necessary to keep the protein in a mixture of detergent (*n*-dodecyl- β -D-maltopyranoside or *n*-decyl- β -D-maltopyranoside) and lipids {1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine, and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-[phosphorac-(1-glycerol)]}; to maintain a strongly re-

Howard Hughes Medical Institute, Laboratory of Molecular Neurobiology and Biophysics, Rockefeller University, 1230 York Avenue, New York, NY 10021, USA.

*To whom correspondence should be addressed. E-mail: mackinn@rockefeller.edu

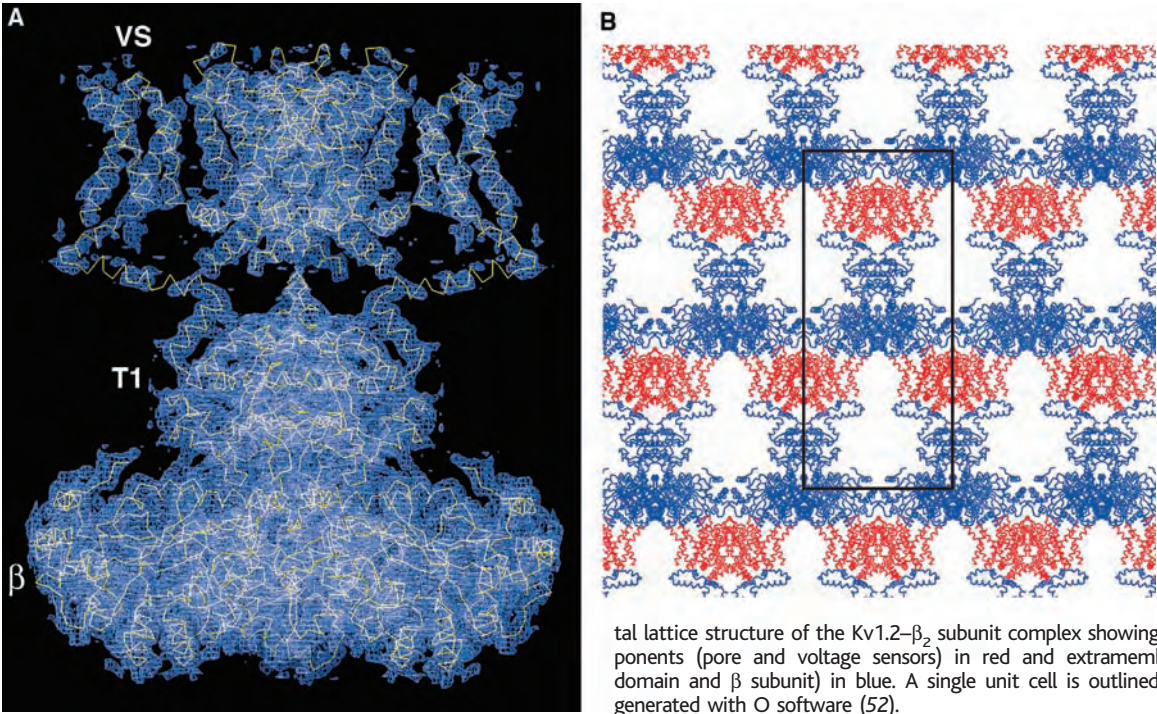


Fig. 1. Electron density and crystal lattice of the Kv1.2-β2 subunit complex. (A) The electron density map (blue mesh) was calculated at 2.9 Å resolution using $2F_o - F_c$ (where F_o is the observed structure factor and F_c is the calculated structure factor) amplitudes and phases from a model without voltage sensors and T1-S1 linkers. It is contoured at 0.8σ . A C_α trace of the final model is shown (yellow). The integral membrane pore and voltage sensors (VS) are on top, the T1 domain in the middle, and the β subunit on the bottom. (B) Crystal lattice structure of the Kv1.2-β2 subunit complex showing integral membrane components (pore and voltage sensors) in red and extramembranous components (T1 domain and β subunit) in blue. A single unit cell is outlined in black. The figure was generated with O software (52).

ducing environment with dithiothreitol (DTT) and tri(2-carboxyethyl)phosphine hydrochloride (TCEP); and to keep the protein in an oxygen-depleted atmosphere. Crystals in space group *I*4, grown by vapor diffusion and frozen in liquid nitrogen, diffracted x-rays to 2.9 Å Bragg spacings at the synchrotron (Table 1). Phases were determined by molecular replacement by using the crystal structure of the T1 domain-β2 complex (PDB ID 1EXB) (12). Electron density for the transmembrane pore and voltage sensors visible in initial maps (Fig. 1A) allowed us to construct a model through cycles of model building, refinement, and map calculation. Figure 1B shows the structure of the crystal lattice, which consists of layers of membrane-spanning regions (pore and voltage sensors in red) alternating with extramembranous regions (T1 domains and β subunits in blue). This arrangement closely mimics a native membrane organization with coplanar arrays of transmembrane elements pointed in the same direction. The quality of electron density varies over the molecule, with T1 and β stronger (mean *B* factor = 59 Å²), and the pore (mean *B* factor = 159 Å²) and voltage sensor (mean *B* factor = 162 Å²) weaker. This variation is understandable in terms of lattice contacts within the crystal and in terms of the function of different components of the channel. The T1 domains and β subunits are rigid structures and well packed in the β subunit layer. The pore is somewhat more flexible in the inner leaflet of the membrane, and the voltage sensors are highly mobile domains that are suspended, making few contacts with neighboring protein molecules. The β subunit, T1 domain, pore (S5, pore helix,

Table 1. Data collection, phasing, and refinement statistics. *I*, intensity; *F*, amplitude; ASU, asymmetric unit; NSLS, National Synchrotron Light Source.

Data collection (<i>I</i> ≥ −3σ):	
Space group	<i>I</i> 4
Cell parameters (Å)	113.6, 113.6, 260.5
Source	NSLS X-25
Resolution (Å)	30.0 to 2.9
Wavelength (Å)	1.1
No. of reflections (total)	105,594
No. of reflections (unique)	36,056
Completeness (%)	99.0 (99.9)*
<i>R</i> _{sym} (%)†	7.2 (47.7)
<i>I</i> /σ	10.8 (2.1)
Refinement (30 to 2.9 Å, <i>F</i> ≥ 0σ _F):	
No. of reflections (work + free)	33,343
No. of molecules/ASU	No. of atoms 4,997
<i>R</i> _{work} / <i>R</i> _{free} (%)‡	1
Bond length/angle§	22.2/25.2
Ramachandran	0.01 Å/1.3°
	99.8% (allowed region)/0.02% (disallowed region)

*The parentheses indicate the outer resolution shell (3.0 to 2.9 Å). † $R_{sym} = \sum |I_i - \langle I_i \rangle| / \sum I_i$, where $\langle I_i \rangle$ is the average intensity of symmetry-equivalent reflections. ‡ R factor = $\sum |F_o - F_c| / \sum F_o$; 5% of the data that were excluded from refinement were used in the *R*_{free} calculation. §The root mean square deviation from ideality for bond angle and length is presented.

filter, and S6), and the S4-S5 linker helix were built as essentially complete models with most side chains included. Four main transmembrane segments of the voltage sensor (S1 to S4) were built without connecting the loops S1-S2, S2-S3, and S3-S4. Prior knowledge of the KvAP structure was helpful in building the voltage sensor, because electron density for side chains was weakest in this region of the channel. The linker connecting the T1 domain to S1 was built as two poly-glycine helices. The complete Kv1.2 channel-β2 subunit complex has fourfold symmetry corresponding to the tetrad axis of the *I*4 unit cell (Fig. 2,

A to C). The dimensions of the tetramer are approximately 135 Å by 95 Å by 95 Å. The pore and four voltage sensors are ~30 Å in length across the membrane, corresponding to the known thickness of the hydrophobic core of the membrane. The T1 domain and β subunit are each ~40 Å in length along the fourfold axis. The T1 domain is offset (parallel to the fourfold axis) from the intracellular pore opening by ~15 to 20 Å. The α-helical T1-S1 linker appears to function as a spacer to maintain separation between the transmembrane pore and intracellular regions of the channel.

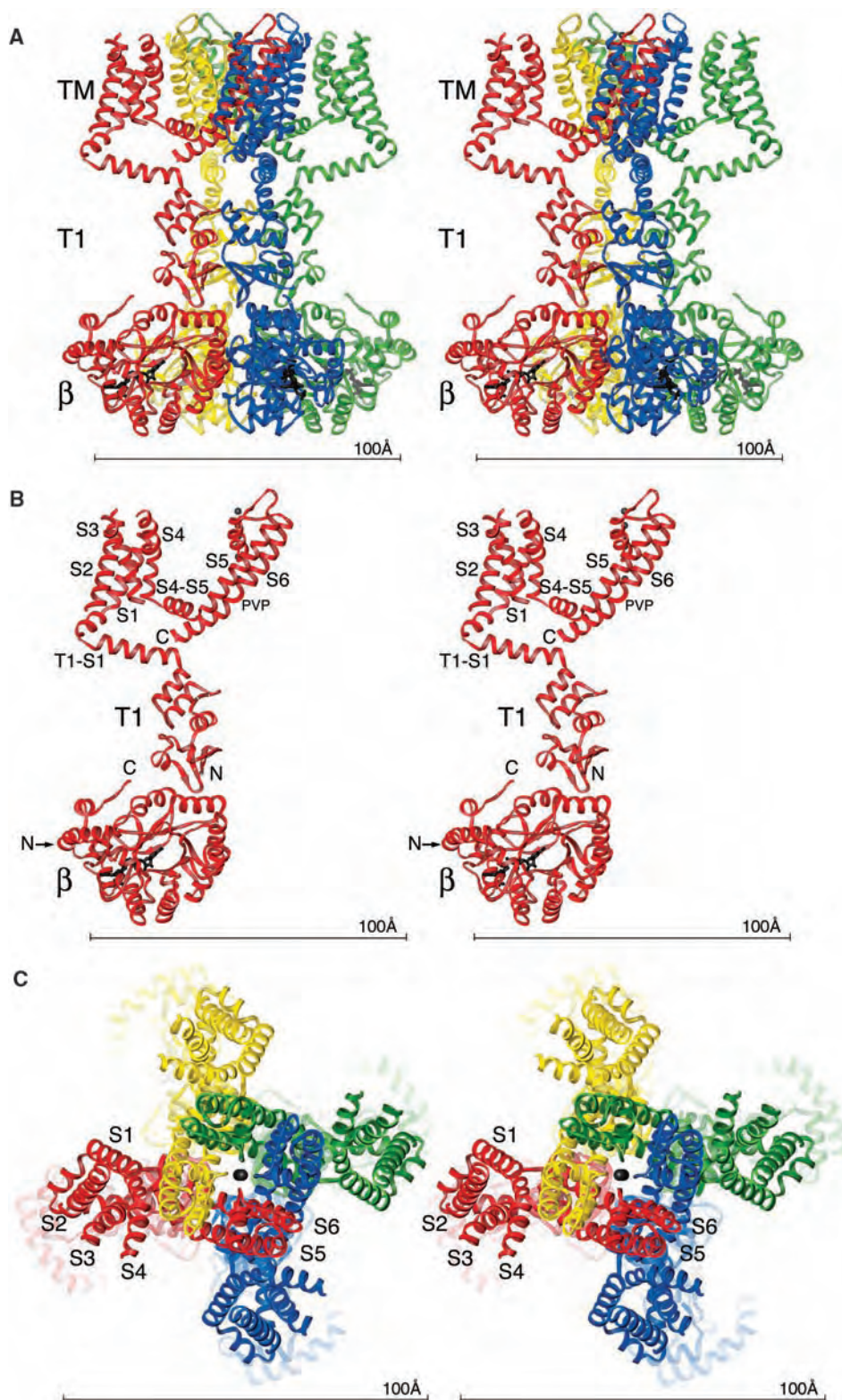


Fig. 2. Views of the Kv1.2- β_2 subunit complex. (A) Stereoview of a ribbon representation from the side, with the extracellular solution above and the intracellular solution below. Four subunits of the channel (including the T1 domain, voltage sensor, and pore) are colored uniquely. Each subunit of the β subunit tetramer is colored according to the channel subunit it contacts. The NADP^+ cofactor bound to each β subunit is drawn as black sticks. TM indicates the integral membrane component of the complex. (B) Stereoview of a single subunit of the channel and β subunit viewed from the side. Labels correspond to transmembrane helices (S1 to S6); the Pro-Val-Pro sequence in S6 (PVP); and the N (N) and C (C) termini of the Kv1.2 and β subunits. The position of the N terminus of the β subunit, which is located on the side furthest away from the viewer, is indicated by an arrow. (C) Stereoview of a ribbon representation viewed from the extracellular side of the pore. Four subunits are colored uniquely. The figure was generated with Ribbons (53).

The α helices of the ion conduction pore (S5, pore helix, and S6) (Fig. 2B) relate to the α helices of the voltage sensor in an unexpected manner. The voltage sensor is latched around the pore of an adjacent subunit. This is most easily appreciated by viewing the channel along its fourfold axis from outside the cell

and noting the manner in which the pore and voltage-sensor units engage each other around the tetramer (Fig. 2C). The connection between the pore and the voltage sensor of a given subunit is made by the S4-S5 linker helix, which runs parallel to the intracellular membrane surface just at the level of the inner-

helix bundle crossing (Fig. 2, A and B). The inner-helix bundle was defined in the KcsA (a non-voltage-dependent K^+ channel) channel as the right-handed bundle of four inner helices (one from each subunit, S6 in Kv1.2) that line the pore on the intracellular side of the selectivity filter (Fig. 3, A and B) (4). This

bundle structure is important because it forms an expandable constriction (the bundle crossing) for opening and closing the ion-conduction pore, known as the activation gate.

In the Kv1.2- $\beta 2$ crystal structure, the voltage-sensor helices are somewhat tilted in the membrane, but they have a native-like transmembrane orientation with hydrophilic loops between helices at the membrane surfaces. In the accompanying Research Article (35), which is focused on voltage-dependent gating, we outline the evidence that the connections between the voltage sensors and pore have maintained a native configuration.

The ion conduction pore. The structures of four different prokaryotic K^+ channels have been determined by x-ray crystallography, including KcsA (4, 36), MthK (37) (a Ca^{2+} -gated K^+ channel), KirBac (38) (an inward-rectifier K^+ channel), and KvAP (30) (a voltage-dependent K^+ channel). A comparison of these channel structures emphasizes two

important aspects of K^+ channels: (i) the structure of the K^+ selectivity filter on the extracellular side of the pore is highly conserved; and (ii) the inner pore (between the selectivity filter and intracellular solution) varies in its conformation. Structural conservation of the selectivity filter underlies the conserved mechanism of selective ion conduction in K^+ channels, whereas variation of the inner pore structure is related to conformational changes that open and close the pore. In the published structures, KcsA and KirBac appear to be closed on the basis of the dimensions of the pore at the narrowest point of the inner-helix bundle (Fig. 3, A and B, the bundle crossing), whereas MthK and KvAP appear to be open. In Fig. 3, A and B, two subunits of the Kv1.2 pore (red) have been superimposed on KcsA (closed, gray) and KvAP (opened, blue) (KvAP inner helices are in a similar position to MthK inner helices). The inner helices of Kv1.2 are in between KcsA and KvAP but much closer

to KvAP at the inner-helix bundle crossing. The pore of Kv1.2 has a diameter of about 12 Å at the bundle crossing, suggesting that the pore in the crystal structure is open.

The superposition of pores in Fig. 3 provides a vivid picture of the large inner-helical motions that must occur when a K^+ channel opens. Functional studies in *Shaker* K^+ channels are consistent with this picture (39, 40). The accessibility of site-directed cysteine residues introduced into the inner helices points to the bundle crossing as the functional constriction for closure. Positions “below” the bundle crossing react with sulfhydryl reagents added to the intracellular side of the membrane whether or not the channel is open, but positions “above” the bundle crossing require the channel to open (40). A tight constriction prevents even Ag^+ ions from passing the bundle crossing when the *Shaker* K^+ channel’s pore is closed, but after it opens, large organic quaternary ammonium cations and hydrophobic polypeptides can enter into the inner pore and plug Kv channels (15, 16, 41, 42).

The inner helices of Kv1.2 contain the amino acid sequence Pro-Val-Pro, which is highly conserved among *Shaker* family Kv channels. These amino acids curve the S6 inner helices so that they run almost parallel to the membrane near the intracellular surface (Fig. 2B and Fig. 3, A and B). Curved inner helices appear to be a specialized adaptation for Kv channels, not *Shaker* channels per se. In KvAP, a prokaryotic Kv channel, a Gly residue produces a similar curvature of the inner helices (Fig. 3, A and B). Thus, the sequence Pro-X-Pro is one way to produce the curve and glycine is another. Both achieve the same structural requirement. In (35), we show why a curved inner helix is required for coupling the voltage sensor to the pore in Kv channels.

The α -helical linkers connecting the T1 domain to S1 project in a radial direction outward from the channel’s central axis to create a wide space between the T1 domain and the pore (Fig. 2A). This would seem to accomplish two functions: It would allow the inner helices to undergo their large gating movements to open the pore without interference from the T1 domain or linkers, and it would provide low-resistance diffusion pathways between the pore entryway and the cytoplasm. A surface rendering of Kv1.2 shows one of these diffusion pathways, termed a side portal, above the T1 domain (Fig. 4A). Four side portals each have a diameter of 15 to 20 Å. The electrostatic potential of this region will attract cations owing to the presence of multiple negatively charged amino acids on the surface (Fig. 4A, red), particularly on the rim of the side portals. This feature of the channel is reminiscent of the intracellular entryway to the acetylcholine receptor ion channel (43).

The dimensions and electrostatic properties of the side portals provide information on an

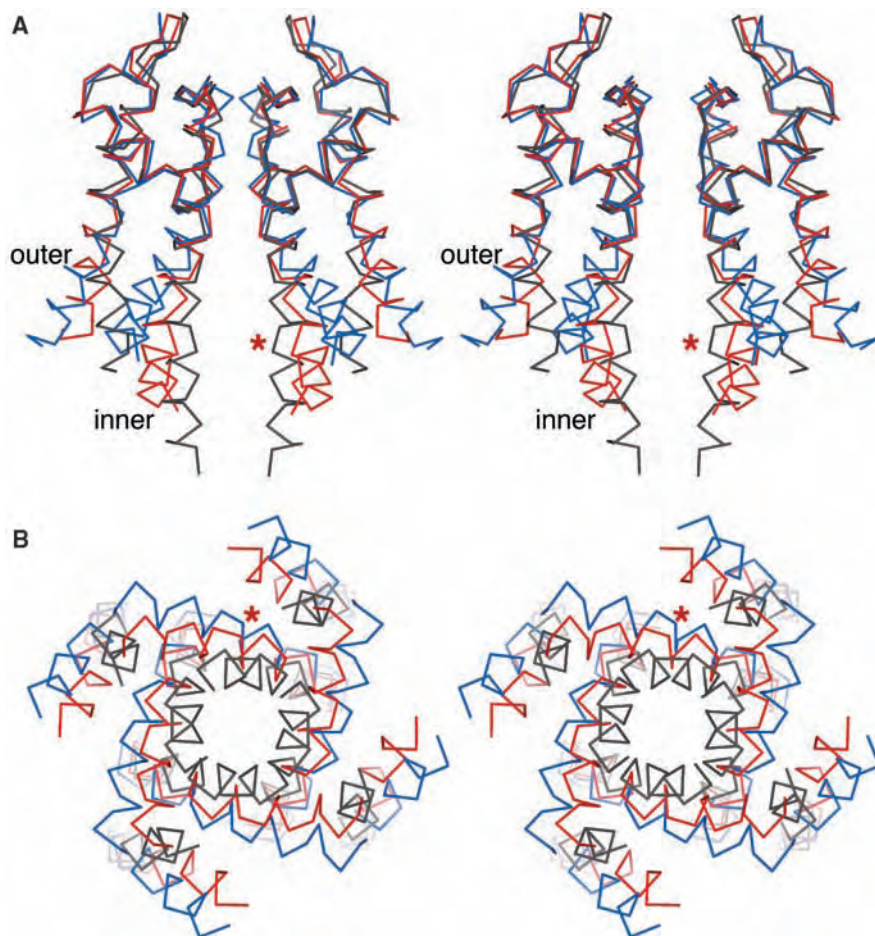


Fig. 3. Comparison of the pore of three K^+ channels. Stereoviews of the Kv1.2 K^+ channel (red, residues 325 to 418), the KcsA K^+ channel (gray, PDB ID 1K4C, residues 26 to 119), and the KvAP K^+ channel (blue, PDB ID 1ORQ, residues 147 to 240) are shown in (A), with two subunits viewed from the side and in (B), with four subunits viewed along the pore axis from the intracellular solution. Channels were superimposed by aligning main-chain atoms of the selectivity filter and pore helices. Outer and inner helices refer to S5 and S6 in Kv1.2. The inner helices form the bundle crossing at the narrowest point, as shown in (B). An asterisk indicates the position of the Pro-Val-Pro sequence on the inner (S6) helix of Kv1.2 and also corresponds to the position of Gly 229 in KvAP. The figure was generated with Molscript (54).

important form of gating known as N-type or “ball and chain” inactivation that occurs in a subset of *Shaker* family Kv channels (15, 16). After membrane depolarization, this subset produces transient K⁺ currents called “A currents,” rather than sustained “delayed rectifier” K⁺ currents (1). A-current channels and delayed-rectifier channels confer different electrical behaviors on cells. For example, delayed-rectifier channels such as Kv1.2 return the membrane potential promptly to its resting value after an action potential, whereas A-current channels allow certain neurons to fire action potentials at a frequency proportional to the total synaptic input current from other neurons (44). From a structural perspective, the distinction between A-current channels and delayed rectifiers is subtle: A-current channels have just the right composition of amino acids on their intracellular N terminus—or on the N terminus of an associated β subunit—to allow the terminus itself to enter the inner pore and plug (inactivate) the Kv channel after it opens its activation gate (15, 16, 32). This kind of N terminus, called an inactivation gate or inactivation peptide, makes an A-current Kv channel transient. Upon membrane depolarization, the pore opens and conducts K⁺ ions, but shortly thereafter it becomes inactivated, as demonstrated by Aldrich and colleagues (15, 16). These authors also defined the amino acids necessary for the N terminus to inactivate the pore of a *Shaker* family Kv channel (45, 46). The first ~10 amino acids should be predominantly hydrophobic, followed by a sequence of hydrophilic and positively charged amino acids. They concluded that the hydrophobic stretch becomes buried in a hydrophobic environment and that the hydrophilic region is important for long-range electrostatic interactions between the inactivation gate and pore. Further studies using the mammalian homolog of *Shaker*, Kv1.4, led to the proposal that the hydrophobic region of the N terminus (of a β subunit inactivation peptide) reaches into the inner pore, which is lined by hydrophobic amino acids, and that positively charged amino acids from the hydrophilic region make electrostatic interactions with negatively charged amino acids from the T1 domain and linkers to S1 (12, 47). We can now see in the Kv1.2 structure that the corresponding negatively charged amino acids reside on the surface of the side portals (Fig. 4A, asterisk). It is easy to imagine an inactivation peptide in an A-current Kv channel snaking through a side portal with its positive amino acids lying against the negative surface and its hydrophobic amino acids extending to the inner pore (Fig. 4B).

Relationship between the pore, T1 domain, and β subunit catalytic site. The first and, to date, the only firmly established role of β subunits of *Shaker* Kv channels is to confer inactivation (32, 48). The β_1 subunit has on its N terminus an amino acid sequence that

has the chemical requirements described above for an inactivation peptide. Thus, when β_1 is coexpressed with a *Shaker* Kv channel, it produces N-type inactivation (32). The N terminus of the structured region of the β_2 subunit in the Kv1.2- β_2 complex is located roughly 80 Å from the pore, whereas the N terminus of the structured region of the channel (T1) is about 50 Å from the pore (Fig. 2B). *Shaker* Kv channels that possess N-type inactivation peptides have at least 30 additional amino acids preceding the T1 domain, and the β_1 subunit has approximately 70 additional amino acids on

its N terminus. As extended polypeptide chains, these N termini are easily long enough to reach the pore to produce N-type inactivation.

The β subunits probably have other functions besides inactivation, because not all β subunits have an N terminal inactivation peptide. For example, the β_2 subunit used to produce the Kv1.2- β_2 crystals, although nearly identical to β_1 through the structured region of β_2 (amino acid 36 to 361), does not have an N terminal inactivation sequence and does not produce inactivation when associated with Kv channels in membranes (48, 49). The β subunits must

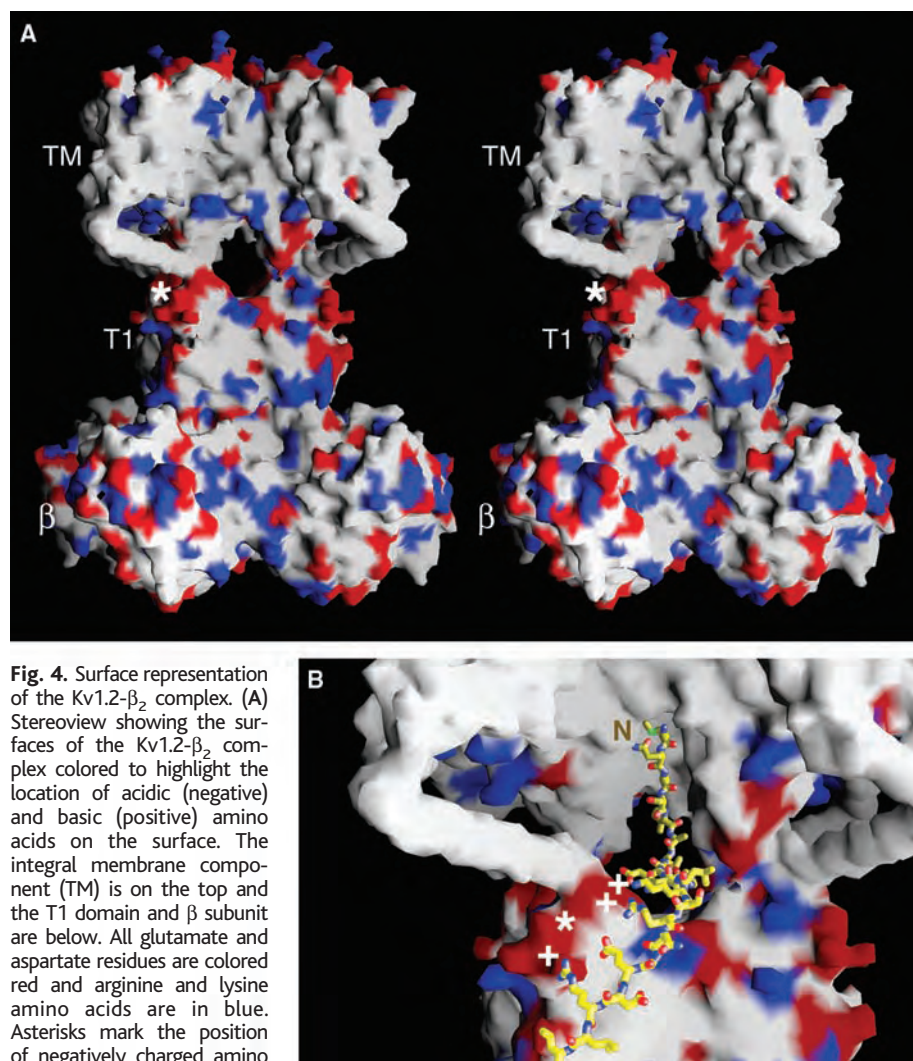


Fig. 4. Surface representation of the Kv1.2- β_2 complex. (A) Stereoview showing the surfaces of the Kv1.2- β_2 complex colored to highlight the location of acidic (negative) and basic (positive) amino acids on the surface. The integral membrane component (TM) is on the top and the T1 domain and β subunit are below. All glutamate and aspartate residues are colored red and arginine and lysine amino acids are in blue. Asterisks mark the position of negatively charged amino acids Glu¹²⁸, Asp¹²⁹, and Glu¹³⁰ on T1, which, in the Kv1.4 K⁺ channel, interact with positively charged amino acids on inactivation peptides (47). The large holes (side portals) above the T1 domain and below the integral membrane channel are passages connecting the cytoplasm to the pore. (B) Hypothetical model for the binding of an inactivation peptide to the channel. A portion of the surface of the T1 and integral membrane components of the complex is shown in the same orientation and colored as in (A). An inactivation peptide corresponding to the 22 N-terminal residues of β_1 has been modeled in an extended conformation to enter a side portal and plug the inner pore of the channel. Three positive residues (+) located at positions 13, 15, and 20 from the N terminus were placed at the site of interaction with the negatively charged patch of residues on T1 [asterisk (12)], whereas the hydrophobic residues at the N terminus were placed at the inner pore where they are known to bind and block the flow of potassium (47). To show entry of the peptide's N terminus into the inner pore in this figure, the following portions of the channel subunit closest to the viewer were removed: residues 312 to 329 (S4-S5 linker and part of S5) and residues 405 to 421 (the C-terminal end of S6). The figure was generated with GRASP software (55).

have additional functions. The observation that β subunits can influence levels of channel expression has led to one hypothesis that they may function as chaperones for the channel (50, 51). The presence of conserved catalytic residues for hydride transfer in the enzyme active site of β subunits forms the basis of another hypothesis, that β subunits are important for a catalytic function (22–24). The Kv channel could regulate the activity of the β subunit, or perhaps more likely, the β subunit could regulate the activity of the channel. A β subunit “redox sensor” would allow direct coupling of membrane electrical activity to the redox state of a cell.

The active site of β subunits has an NADP⁺ cofactor and catalytic residues for hydride transfer. This cofactor is present in the Kv1.2- β 2 crystal structure (Fig. 5A). In addition, there is poorly defined extra density overlying the active site in all of our electron density maps (Fig. 5B). The surface of the active-site cleft, where this material is bound, is quite hydrophobic and has many aromatic amino acids (Fig. 5B). The density is probably either from a polypeptide chain or from a large organic molecule. If it represents a polypeptide chain, it could either be part of the channel or it could be co-isolated with the channel during purification. The channel in the crystal structure contains 31 disordered amino acids on its N terminus (N terminal to T1) and 78 on its C terminus. Binding either the N terminus or the C terminus near the β subunit active site, with an affinity that depends on whether the cofactor is oxidized or reduced, could affect Kv channel function. For example, in some *Shaker* family Kv channels, binding of the polypeptide chain preceding T1 could affect the availability of the inactivation gate. The C-terminal polypeptide could potentially influence gating because it is connected directly to the inner

helices, which must undergo large movements to gate the pore open and closed. The Kv1.2- β 2 crystal structure should help in the design of experiments to test these ideas.

Summary. This work describes the crystal structure of the mammalian *Shaker* family Kv channel Kv1.2 in complex with the β ₂ subunit. Both proteins were coexpressed in *Pichia pastoris*. Two critical factors were essential for obtaining crystals and determining the structure. A mixture of lipids and detergent was used throughout purification and crystallization, and many steps were taken to minimize oxidation. The importance of lipids in this project may suggest the general application of lipids in membrane protein structural studies in the future. The I4 crystal lattice of alternating “membrane” and water layers comes very close to a native arrangement for the channel. The relationships between the integral membrane components, the intracellular T1 domain, and the β subunit bound to the T1 domain are consistent with electrophysiological studies of the inactivation process in *Shaker* K⁺ channels. The pore of Kv1.2 is similar in structure to that of prokaryotic K⁺ channels. Curved inner helices are produced by the sequence Pro-X-Pro in *Shaker* family K⁺ channels and by Gly in KvAP and probably many other Kv channels. The Kv1.2 pore is in an open conformation. Large side portals between the T1 domain and the pore connect the pore to the cytoplasm. Electrostatic properties of the side portals are consistent with long-range electrostatic steering of inactivation peptides. An NADP⁺ cofactor, catalytic residues, and extra density at the β subunit active-site cleft suggest experiments for testing the influence of β subunit catalytic function on K⁺ channel activity. The voltage sensors are in the “membrane” layer of the crystal and appear to be unperturbed.

References and Notes

- B. Hille, *Ion Channels of Excitable Membranes* (Sinauer Associates, Sunderland, MA, ed. 3, 2001).
- G. Yellen, *Q. Rev. Biophys.* **31**, 239 (1998).
- R. MacKinnon, *Angew. Chem. Int. Ed. Engl.* **43**, 4265 (2004).
- D. A. Doyle et al., *Science* **280**, 69 (1998).
- D. H. Hackos, T. H. Chang, K. J. Swartz, *J. Gen. Physiol.* **119**, 521 (2002).
- Z. Lu, A. M. Klem, Y. Ramu, *J. Gen. Physiol.* **120**, 663 (2002).
- S. M. Webster, D. del Camino, J. P. Dekker, G. Yellen, *Nature* **428**, 864 (2004).
- B. L. Tempel, D. M. Papazian, T. L. Schwarz, L. Y. Jan, Y. N. Jan, *Science* **237**, 770 (1987).
- A. Kamb, L. E. Iverson, M. A. Tanouye, *Cell* **50**, 405 (1987).
- O. Pongs et al., *EMBO J.* **7**, 1087 (1988).
- A. Kreusch, P. J. Pfaffinger, C. F. Stevens, S. Choe, *Nature* **392**, 945 (1998).
- J. M. Gulbis, M. Zhou, S. Mann, R. MacKinnon, *Science* **289**, 123 (2000).
- O. Sokolova, L. Kolmakova-Partensky, N. Grigorieff, *Structure (Camb.)* **9**, 215 (2001).
- W. R. Kobertz, C. Williams, C. Miller, *Biochemistry* **39**, 10347 (2000).
- T. Hoshi, W. N. Zagotta, R. W. Aldrich, *Science* **250**, 533 (1990).
- W. N. Zagotta, T. Hoshi, R. W. Aldrich, *Science* **250**, 568 (1990).
- S. Sewing, J. Roeper, O. Pongs, *Neuron* **16**, 455 (1996).
- W. Yu, J. Xu, M. Li, *Neuron* **16**, 441 (1996).
- H. Rehm, M. Lazdunski, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4919 (1988).
- D. N. Parcej, V. E. Scott, J. O. Dolly, *Biochemistry* **31**, 11084 (1992).
- V. E. Scott et al., *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1637 (1994).
- T. McCormack, K. McCormack, *Cell* **79**, 1133 (1994).
- S. W. Chouinard, G. F. Wilson, A. K. Schlimgen, B. Ganetzky, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6763 (1995).
- J. M. Gulbis, S. Mann, R. MacKinnon, *Cell* **97**, 943 (1999).
- K. Nakahira, G. Shi, K. J. Rhodes, J. S. Trimmer, *J. Biol. Chem.* **271**, 7084 (1996).
- F. J. Sigworth, *Q. Rev. Biophys.* **27**, 1 (1994).
- F. Bezanailla, *Physiol. Rev.* **80**, 555 (2000).
- F. Sesti, S. Rajan, R. Gonzalez-Coloso, N. Nikolaeva, S. A. Goldstein, *Nat. Neurosci.* **6**, 353 (2003).
- V. Ruta, Y. Jiang, A. Lee, J. Chen, R. MacKinnon, *Nature* **422**, 180 (2003).
- Y. Jiang et al., *Nature* **423**, 33 (2003).
- W. Stuhmer et al., *EMBO J.* **8**, 3235 (1989).
- J. Rettig et al., *Nature* **369**, 289 (1994).
- D. N. Parcej, L. Eckhardt-Strelau, *J. Mol. Biol.* **333**, 103 (2003).
- Materials and methods are available as supporting material on Science Online.
- S. B. Long, E. B. Campbell, R. MacKinnon, *Science* **309**, 903 (2005); published online 7 July 2005 (10.1126/science.1116270).
- Y. Zhou, J. H. Morais-Cabral, A. Kaufman, R. MacKinnon, *Nature* **414**, 43 (2001).
- Y. Jiang et al., *Nature* **417**, 515 (2002).
- A. Kuo et al., *Science* **300**, 1922 (2003).
- D. del Camino, M. Holmgren, Y. Liu, G. Yellen, *Nature* **403**, 321 (2000).
- Y. Liu, M. Holmgren, M. E. Jurman, G. Yellen, *Neuron* **19**, 175 (1997).
- D. del Camino, G. Yellen, *Neuron* **32**, 649 (2001).
- C. M. Armstrong, *J. Gen. Physiol.* **58**, 413 (1971).
- A. Miyazawa, Y. Fujiyoshi, M. Stowell, N. Unwin, *J. Mol. Biol.* **288**, 765 (1999).
- J. A. Connor, C. F. Stevens, *J. Physiol.* **213**, 1 (1971).
- R. D. Murrell-Langnado, R. W. Aldrich, *J. Gen. Physiol.* **102**, 949 (1993).
- R. D. Murrell-Langnado, R. W. Aldrich, *J. Gen. Physiol.* **102**, 977 (1993).
- M. Zhou, J. H. Morais-Cabral, S. Mann, R. MacKinnon, *Nature* **411**, 657 (2001).
- S. H. Heinemann, J. Rettig, H. R. Graack, O. Pongs, *J. Physiol.* **493**, 625 (1996).

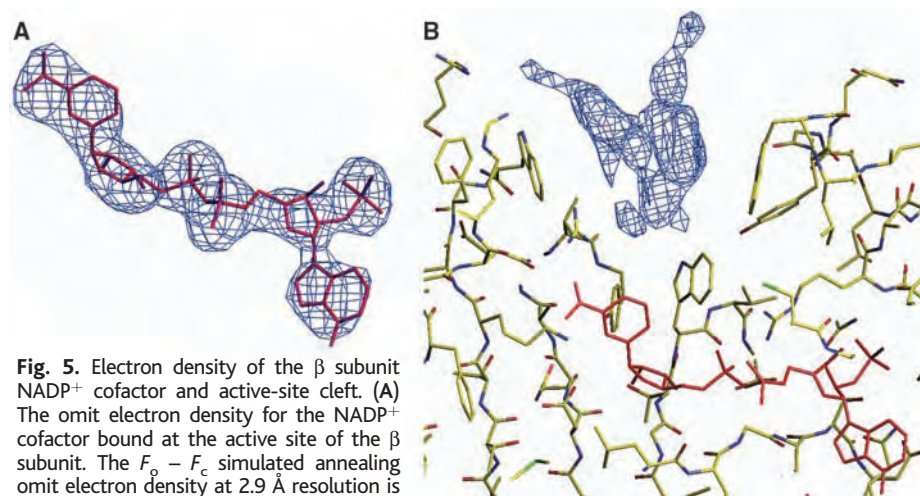


Fig. 5. Electron density of the β subunit NADP⁺ cofactor and active-site cleft. (A) The omit electron density for the NADP⁺ cofactor bound at the active site of the β subunit. The $F_o - F_c$ simulated annealing omit electron density at 2.9 Å resolution is contoured at 4 σ (blue mesh) and shown with the NADP⁺ molecule drawn as red sticks. (B) “Extra” density in the active-site cleft of the β subunit. A $2F_o - F_c$ electron density map to 2.9 Å resolution and contoured at 0.8 σ is drawn in a pocket above the catalytic residues. The NADP⁺ cofactor is shown as red sticks.

49. K. McCormack, T. McCormack, M. Tanouye, B. Rudy, W. Stuhmer, *FEBS Lett.* **370**, 32 (1995).
50. G. Shi et al., *Neuron* **16**, 843 (1996).
51. N. Nagaya, D. M. Papazian, *J. Biol. Chem.* **272**, 3022 (1997).
52. T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr.* **A47**, 110 (1991).
53. M. Carson, *Methods Enzymol.* **277**, 493 (1997).
54. P. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).
55. A. Nicholls, K. A. Sharp, B. Honig, *Proteins* **11**, 281 (1991).
56. We thank A. Lee, V. Ruta, and members of the MacKinnon

laboratory for helpful discussions; R. Jain for initial experiments with lipids; Q. Wang and B. T. Chait for mass spectrometry; R. Dutzler for assistance with data collection; O. Pongs for Kv1.2 DNA; J. Trimmer for β_2 subunit DNA; Brookhaven National Laboratory (National Synchrotron Light Source beamlines X25 and X29) and the Swiss Light Source (beamline PX1) staff for assistance in data collection; and W. Chin for help with manuscript preparation. This work was supported in part by NIH grant no. GM43949 to R.M. and NIH grant no. RR00862 to B. T. Chait. R.M. is an Investigator in the Howard Hughes Medical Institute. Atomic

coordinates and structure factors have been deposited with the Protein Data Bank with accession ID 2A79.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1116269/DC1
Materials and Methods
References

17 June 2005; accepted 5 July 2005

Published online 7 July 2005;

10.1126/science.1116269

Include this information when citing this paper.

Voltage Sensor of Kv1.2: Structural Basis of Electromechanical Coupling

Stephen B. Long, Ernest B. Campbell, Roderick MacKinnon*

Voltage-dependent ion channels contain voltage sensors that allow them to switch between nonconductive and conductive states over the narrow range of a few hundredths of a volt. We investigated the mechanism by which these channels sense cell membrane voltage by determining the x-ray crystal structure of a mammalian *Shaker* family potassium ion (K^+) channel. The voltage-dependent K^+ channel Kv1.2 grew three-dimensional crystals, with an internal arrangement that left the voltage sensors in an apparently native conformation, allowing us to reach three important conclusions. First, the voltage sensors are essentially independent domains inside the membrane. Second, they perform mechanical work on the pore through the S4-S5 linker helices, which are positioned to constrict or dilate the S6 inner helices of the pore. Third, in the open conformation, two of the four conserved Arg residues on S4 are on a lipid-facing surface and two are buried in the voltage sensor. The structure offers a simple picture of how membrane voltage influences the open probability of the channel.

Voltage-dependent ion channels open in response to changes in voltage across the cell membrane (1). In this process, the membrane electric field performs mechanical work to alter the channel's conformation within the membrane. The work arises from the force exerted by the electric field on charged amino acids, termed gating charges (1–3). The size of the gating charge is very large (4), accounting for the exquisite sensitivity of voltage-dependent ion channels to small changes in membrane voltage. To understand this process, one must first answer two questions: How do gating charges move within the membrane electric field? And how are these movements mechanically coupled to opening and closing of the pore?

No experimentally based model has yet provided answers to both of these questions. So far, little progress has been made toward the second question concerning the mechanical coupling of voltage-sensor movements to the pore. Most effort has focused on how the

gating charges move; the main subject of study has been the *Shaker* voltage-dependent K^+ (Kv) channel, and numerous models have been put forth. One fundamental constraint for any model is that when a *Shaker* K^+ channel opens, it transfers the net equivalent of 12 to 14 positive elementary charges across the membrane electric field from inside to outside, and most of this charge is carried by four S4 Arg residues on each of four identical channel subunits (4–6).

A guiding assumption underlying most models for the voltage sensor has been that the S4 helix with its Arg residues is completely (7–10), or mostly (11), sequestered from the membrane, in order to protect the charges from the lipid's low dielectric environment. To accomplish this, most models postulate that the S4 helix inserts into a groove at the interface between adjacent subunits of the K^+ channel tetramer, such that pore α helices S5 and S6 form a wall on one side of S4 and voltage-sensor α helices S1 to S3 form a wall on the other side, the lipid-facing perimeter, to create a gating channel or protein-lined canaliculus for S4 (7–10, 12). This arrangement would allow the S4 helix to move its charged amino acids across the membrane without exposing them to the lipid environment.

How and to what extent S4 moves has been the subject of much debate. Gating-dependent reactivity of sulfhydryl reagents with cysteine residues led to an initial hypothesis of large (~ 15 Å) translations of S4 in some models (13–17). But very small distance changes measured in fluorescence resonance energy transfer (FRET) experiments suggested much smaller movements of S4 across the membrane (18, 19). Crevices surrounding S4 were invoked to account for the sulfhydryl reactivity in the setting of these smaller movements, with translations and/or rotations of S4 occurring across a narrow neck inside an hourglass-shaped canaliculus. In the transporter model hypothesis, the S4 does not change its depth in the membrane at all (less than 3 Å movement) (18, 20). Instead, the field is moved over the S4 charges by alternately opening and closing crevices to the internal and external solutions.

The above models of voltage-dependent gating vary in detail, but they have had two essential features in common. First, the S4 helix is sequestered from the lipid membrane [although Larsson and colleagues proposed that a surface of S4 could be exposed to lipid (11)]. Second, the voltage-sensor helices S1 to S4 are packed tightly against the α helices of the pore. In other words, it was reasonably assumed that voltage-dependent ion channels are like conventional α -helical membrane proteins that form a fairly rigid disk of helices in the membrane. A first hint that a rigid disk of helices might not pertain to voltage-dependent ion channels came from the demonstration that the voltage sensor (S1 to S4) from the *Shaker* Kv channel could be spliced onto the pore of KcsA (a non-voltage-dependent K^+ channel) to confer voltage-dependent gating (21). This finding implied that the voltage sensor might be an almost-independent domain, because if it had to form a large interface through helix packing with the pore, the chimera would likely not function.

The first atomic structures of a prokaryotic Kv channel (KvAP) also implied that the voltage sensors are loosely attached to the pore (22). One of these [Protein Data Bank (PDB) ID 1ORS] was of an isolated voltage-sensor domain, which surprisingly could be expressed in the membrane by itself (without the pore). Another, which is a full-length channel structure (PDB ID 1ORQ), showed the voltage-

Howard Hughes Medical Institute, Laboratory of Molecular Neurobiology and Biophysics, Rockefeller University, 1230 York Avenue, New York, NY 10021, USA.

*To whom correspondence should be addressed.
E-mail: mackinn@rockefeller.edu

sensor domains in a non-native conformation, pulled toward the cytoplasmic side of the pore. Not only had the voltage sensors undergone domain-like movements with respect to the pore, but the sensors seemed to have a great deal of internal flexibility. This was unusual behavior for a membrane protein. The Arg-containing S4 helix formed part of a helix-turn-helix structure (termed a voltage-sensor paddle) through its antiparallel relationship to S3b (the C-terminal half of S3). The paddle was proposed to move at the protein-lipid interface with S3b "above" and S4 "below." Experiments with avidin capture of biotin suggested

that some of the Arg residues move more than 15 Å through the thickness of the membrane (23, 24).

The KvAP studies implied that the voltage sensors are highly mobile; that S4 is not in a canaliculus; and possibly that some S4 Arg residues could be exposed to the lipid membrane, which would allow the voltage-sensing apparatus to exploit opposing electrostatic and hydrophobic forces to gate the channel (23, 25). However, a major weakness of the KvAP studies was directly related to the voltage sensors' mobility: Distortions associated with extraction from the membrane left many aspects of the structure uncertain. Most notably, the connections between the voltage sensors and pore were disrupted. The crystals of Kv1.2 have maintained these connections and thus convey more defin-

itive information on the mechanism of voltage-dependent gating.

Voltage-sensor structure: Relationship between Kv1.2 and KvAP. Electron density maps are weak in the region of the voltage sensor, relative to the remainder of the channel. However, the four transmembrane helices (S1 to S4) were easily recognizable, with side-chain density for some, but not all, amino acids (26). Electron density was present for the first (Arg²⁹⁴), second (Arg²⁹⁷), and third (Arg³⁰⁰) Arg residues on S4, and for two Phe residues (Phe³⁰² and Phe³⁰⁵), establishing the correct register of this helix (Fig. 1). The partial model of the voltage sensor contains helices for S1 (19 amino acids), S2 (residues 219 to 243), S3 (21 amino acids), S4 (residues 288 to 311), and the S4-S5 linker (residues 312 to 325). Com-

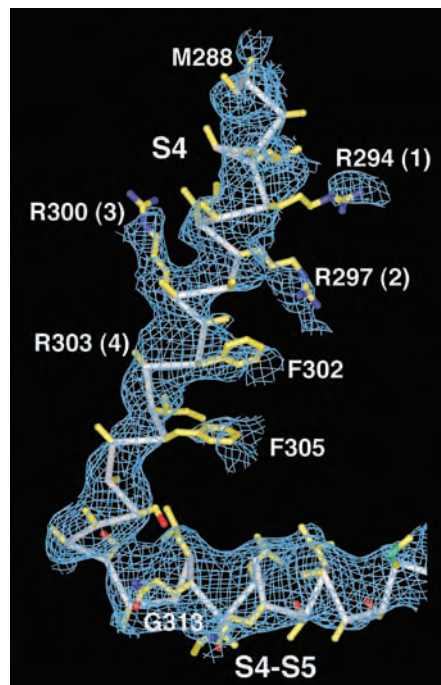


Fig. 1. Density for S4 and the S4-S5 linker calculated without a voltage-sensor model. The electron density map for S4 and the S4-S5 linker (blue mesh) is shown with the final model drawn as a C_{α} trace (gray) with side-chain residues (yellow, blue, and red sticks). Arg (R) residues 1 to 3 on S4 have density for their side chains. Density for these side chains and for two Phe residues (F302 and F305) helped establish the correct register of the S4 helix. The side chain of Arg 4 is truncated after the C_{β} atom (modeled as alanine) because density past this point was not present in the maps. A few other residues in S4 have also been modeled as alanine. Phases for the map were calculated by removing the entire voltage sensor (S1 to S4 helices through residue 313 of the S4-S5 linker) from the model and refining the remaining partial structure of the pore and T1/β complex using a simulated annealing protocol in the CNS software (41). This procedure, which is used to generate a "simulated annealing omit map," essentially eliminates bias in the map. The map is a $2F_o - F_c$ map (where F_o is the observed structure factor and F_c is the calculated structure factor) that was calculated from 30 to 2.9 Å, contoured at 0.5σ , and drawn around the portion of the molecule shown. M, Met; G, Gly.

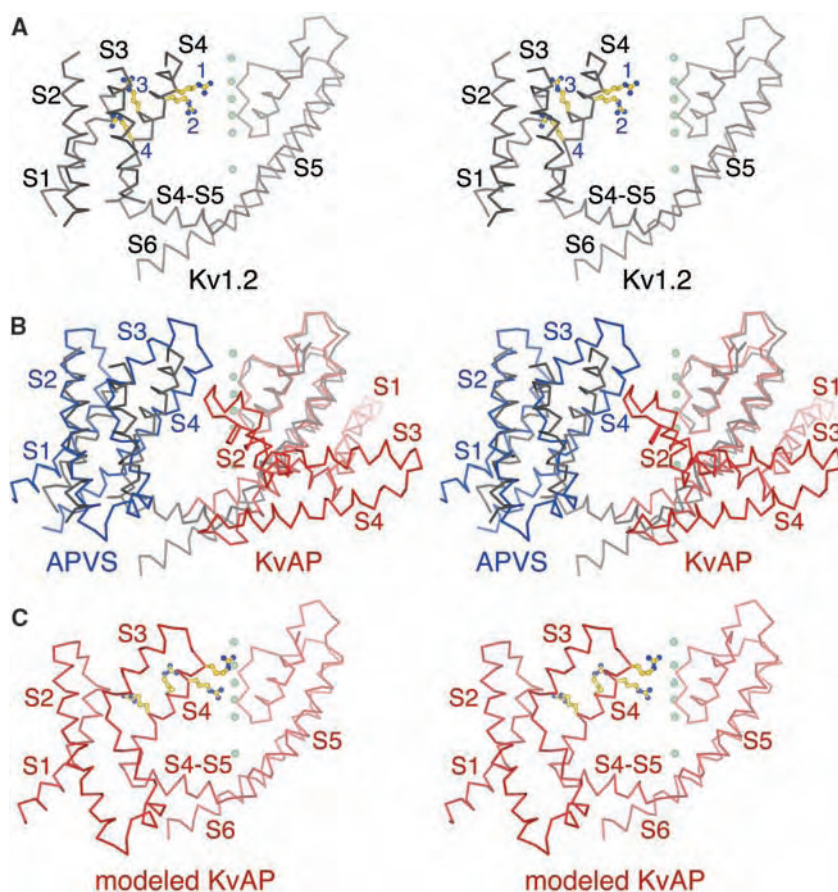


Fig. 2. Stereoviews comparing the Kv1.2 structure with two structures of the prokaryotic Kv channel KvAP. (A) A single subunit of the integral membrane pore and partial model of the voltage sensor of Kv1.2 viewed from the side as a gray C_{α} trace. Arg residues 1 to 4 on the S4 helix (blue labels) are depicted as yellow and blue sticks. The side chain for Arg 4, although not included in the final coordinates, is modeled in a chemically reasonable conformation for the purpose of illustration. (B) The Kv1.2 structure (gray) viewed as in (A) with a full-length crystal structure of KvAP (red C_{α} trace, PDB ID 2A0L) superimposed by alignment of main-chain atoms of the pore helices and selectivity filter, and with an isolated voltage-sensor structure of KvAP (Aeropyrum Pernix Voltage Sensor, or APVS) (blue C_{α} trace, PDB ID 1ORS) (22) superimposed by alignment of main-chain atoms of α helices S1 and S2. (C) A hypothetical model of a single KvAP subunit is shown as a red C_{α} trace with yellow and blue side chains for Arg residues 1 to 4 on the S4 helix. This was constructed by combining the isolated voltage sensor and pore of KvAP according to their positions relative to Kv1.2 as displayed in (B). The S4-S5 linker residues of KvAP (residues 136 to 146) are positioned relative to the pore and voltage sensor based on the Kv1.2 S4-S5 helix. A queue of K^+ ions (green spheres) from the pore are shown as a reference in (A) to (C). The figure was generated with Molscript software (42).

parisons with KvAP crystal structures assisted in the identification of these helices. Most side chains were included on S4, the S4-S5 linker, and S2. S1 and S3 were built with alanine residues. Loops connecting helices S1 to S2 and S3 to S4 were omitted because electron density was weak or absent. The turn connecting S2 to S3, which varies in its conformation in different crystal structures of KvAP (22) (and see PDB ID 2A0L), was also omitted. Thus, this partial model of the voltage sensor is missing several elements, but it still addresses many important questions.

A model of a single subunit of Kv1.2 is shown in Fig. 2A. Two different crystal structures of KvAP superimposed on Kv1.2 show the relationship between these channels (Fig. 2B). A full-length KvAP structure (PDB ID

2A0L, red) is aligned with the pore of Kv1.2, and an isolated voltage-sensor structure of KvAP [PDB ID 1ORS, blue (22)] is aligned with the voltage sensor of Kv1.2. A model of KvAP can be made to look like Kv1.2 by simply repositioning its voltage sensor to more closely resemble the isolated voltage-sensor structure (Fig. 2C). The linker connecting the voltage sensor to the pore in KvAP is the appropriate length and has the correct amphipathicity (see below) to match the linker in Kv1.2. Apparently, upon extraction of KvAP from the lipid membrane, the voltage sensor is dislodged from its proper position. This kind of distortion in multiple crystal structures of KvAP (22) (and see PDB ID 2A0L) probably reflects the importance of a cell membrane to hold the voltage sensor in its proper position. In

Kv1.2, the presence of a T1 domain and its connection to the S1 helix undoubtedly help to maintain a native conformation of the voltage sensor.

We conclude that the basic architecture of the voltage sensor in a membrane is similar in Kv1.2 and KvAP. They both have an antiparallel arrangement of S3 and S4. This arrangement was called a voltage-sensor paddle in KvAP (22). In KvAP, there are two distinct segments of S3, termed S3a and S3b. In Kv1.2, the electron density for S3 appears to be a single helix with a bend, presumably near the connection between S3a and S3b. This finding probably represents a different positioning of the paddle from that which we observed in KvAP. Here we do not distinguish between S3a and S3b in Kv1.2, but simply refer to the antiparallel unit formed by the S3 and S4 helices as the voltage-sensor paddle.

The comparison of Kv1.2 and KvAP serves many useful purposes. First, their fundamental similarity reinforces our confidence in the accuracy of the Kv1.2 voltage-sensor model and shows that S4 and S3 form a voltage-sensor paddle as in KvAP. Second, because Kv1.2 and KvAP are similar, we can consider KvAP functional data in constraining possible motions of the Kv1.2 voltage sensor. Third, certain differences between their structures may provide useful information about movements of the voltage sensor. For example, the voltage-sensor paddles in Kv1.2 are in a slightly different position with respect to the S1 and S2 helices (Fig. 2B). This is understandable if the paddles are mobile, allowing them to move in the gating process. Fourth, S4 (and S3) is nearly two helical turns longer at its extracellular end in Kv1.2 than in KvAP (the Kv1.2 S4 contains two extra helical turns preceding the Arg residues) (compare Fig. 2, A and C). This means the paddle in Kv1.2 will project further into the extracellular solution and therefore may exhibit differences (relative to KvAP) in accessibility to spider toxins and small molecules that interact with the voltage sensor from outside the cell.

Voltage-sensor coupling to the pore:

The S4-S5 linker helix. The S4-S5 linker is an amphipathic α helix that runs parallel to the membrane plane inside the cell, with its hydrophobic surface facing the membrane and its polar surface facing the cytoplasm (Figs. 2A and 3A). The most important aspect of this helix is its position against the pore; it crosses over the top of the S6 inner helix from the same subunit and makes many amino acid contacts with it (Fig. 3, A and B). The S6 inner helix, by curving parallel to the membrane plane, makes a platform or “receptor” for the S4-S5 helix. This allows us to understand why the S6 helix of Kv channels has the sequence Pro-X-Pro, where X is any amino acid (*Shaker* Kv channels), or Gly in the corresponding region (many other Kv channels): to allow the inner helices to curve so they can form the

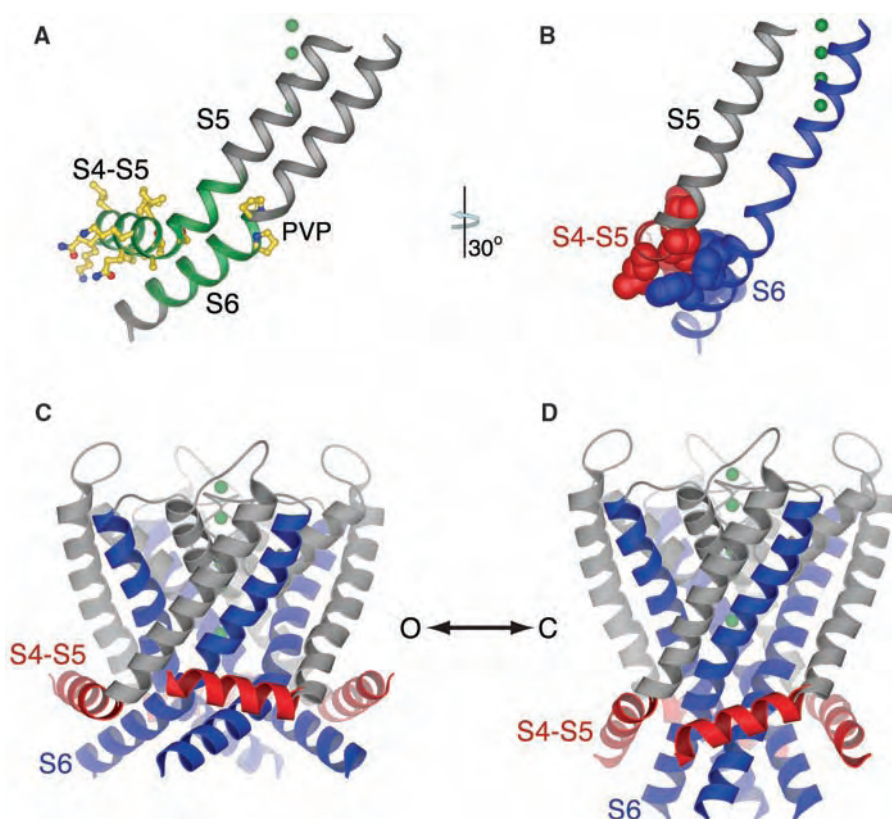


Fig. 3. The connection between the voltage sensor and the pore in the Kv1.2 channel. (A) The S6 inner helix (residues 388 to 421) is shown as a gray and green ribbon with yellow side chains for Pro-Val-Pro (residues 405 to 407), and the S4-S5 linker and S5 from the same subunit (residues 311 to 342) are shown as a gray and green ribbon. Side chains on the S4-S5 linker are yellow (carbon), red (oxygen), and blue (nitrogen). The perspective is from the side of the channel near the intracellular water (below)/membrane (above) interface. Regions colored green were necessary to transfer the *Shaker* voltage sensor to KcsA (32). (B) Residues on the S4-S5 linker in direct contact with residues on S6 are shown as red and blue spheres, respectively. The helices are drawn as ribbons and colored in the following manner: S4-S5, red; S5, gray; and S6, blue. (C) A view of the channel tetramer showing the S4-S5 (red), S5 (gray), and S6 (blue) helices as ribbons. The perspective is from the side of the channel with the extracellular side above and the intracellular side below. (D) Hypothetical model of the Kv1.2 channel with a closed activation gate, showing the S4-S5, S5, and S6 helices colored as in (C). To generate this model, the inner (S6) helices were adjusted from their observed open conformation in (C) to match the inner helices of the KcsA structure (PDB ID 1K4C), which has a closed activation gate. The S4-S5 linkers were then positioned to maintain the interaction with S6 shown in (A) and (B). The transition from an open to closed activation gate results in a downward displacement (toward the intracellular solution) of the amino-terminal end of the S4-S5 linker. A queue of K^+ ions (green spheres) from the pore are shown as a reference in (A) to (D). The figure was generated with Molscript (42).

correct interaction with the S4-S5 linker helix. This interaction is essential for the coupling of voltage-sensor movements to pore opening and closing, which is depicted in Fig. 3, C and D.

Mutations in the Pro-X-Pro sequence and in the S4-S5 linker helix of *Shaker* Kv channels (27–31) have profound effects on gating, which have been described as uncoupling the pore from the voltage sensor. One mutational study leaves little doubt about the correctness and importance of the interaction we see between the S4-S5 linker and the S6 inner helix observed in the Kv1.2 crystal structure. Lu *et al.* characterized the amino acid sequence requirements for engineering voltage dependence into KcsA, an otherwise voltage-independent K⁺ channel (21, 32). They found they had to transfer to KcsA the *Shaker* Kv channel voltage sensor (S1 to S4), the S4-S5 linker, and the C-terminal end of S6. The

required segment of S6 corresponds precisely to the region that makes contact with the S4-S5 linker helix in the Kv1.2 structure (Fig. 3A, in green). Their experiments showed that this interface, formed by the S4-S5 helix against the S6 inner helix, is both necessary and sufficient to reconstruct a functioning voltage sensor on the pore.

Lipid environment of the voltage sensor. The specific interaction between the S4-S5 linker helix and S6 has important consequences for the location of the voltage sensor relative to the pore. Because the linker runs across to the neighboring subunit, the voltage-sensor domains are located at the corners of the square-shaped pore, and they are adjacent to the pore-forming helices of a neighboring subunit (Figs. 2A and 4A). The resulting position of S4 (adjacent to S5 from a neighboring subunit) is in good agreement

with the disulfide cross-bridge studies of Papazian and colleagues (33). Several studies have attempted to determine the distances separating the first Arg residue on the S4 helix from adjacent and diagonal subunits (18, 19). Here we measure these distances (between C_α carbons) to be 45 and 64 Å, respectively. The position of the voltage sensors at the corners of the pore is reminiscent of a model proposed by Sivaprasadarao and colleagues, but in their model, the voltage sensor contacts the pore of its own subunit rather than the pore of the neighboring subunit (34).

The most important consequence of being at the corners of the pore is that the voltage sensors appear to be floating as separate domains from the pore. Aside from the S4-S5 linker interaction with S6, the contacts between a voltage sensor and the pore are not substantial; the tilted S1 helix touches S5 in one place near the extracellular membrane surface, and the S4 helix, which is supposed to move with channel gating, leans against the outer edge of S5 but is not packed tightly against it (Fig. 4, A to C). In a membrane, much of the space separating the hydrophobic surfaces of the pore and the voltage sensor would undoubtedly be filled with lipid molecules (Fig. 4, B and C).

The relative independence of the voltage-sensor domains with respect to the pore in the crystal structure is consistent with several key observations on voltage-sensor function. An independent domain relationship explains why it is possible to transfer a voltage sensor to a non-voltage-dependent K⁺ channel (providing that the complementary surfaces at the linker are satisfied) (21), why the voltage sensors of KvAP can be expressed in isolation (22), and why nature has been able to exploit the S1 to S4 voltage-sensor domain (in the absence of an ion channel pore) to control the activity of a phosphatase enzyme in the cytoplasm (35). The existence of a voltage-dependent phosphatase enzyme is a direct demonstration by nature that a protein wall formed by the pore on one side of S4 is not necessary for the voltage sensor to function. All by itself, this simple arrangement of S1 to S4 helices must be able to undergo a voltage-dependent conformational change in the membrane.

Where are the gating-charge Arg residues on the voltage sensor? Studies of the *Shaker* Kv channel have shown that the first four Arg residues (termed Arg 1 through 4 counting from the extracellular side of S4: residues 294, 297, 300, and 303 in Kv1.2) account for most of the gating charge (5, 6), and these residues are the most conserved among voltage sensors from different Kv channels. The chemical environment of these amino acids on the voltage sensor is a mechanistically important and much-debated issue. In the crystal structure, Arg 1 and 2 are located on the voltage sensor's lipid-facing surface (Fig. 4D and Fig. 5, A and B). The first may be near enough to the mem-

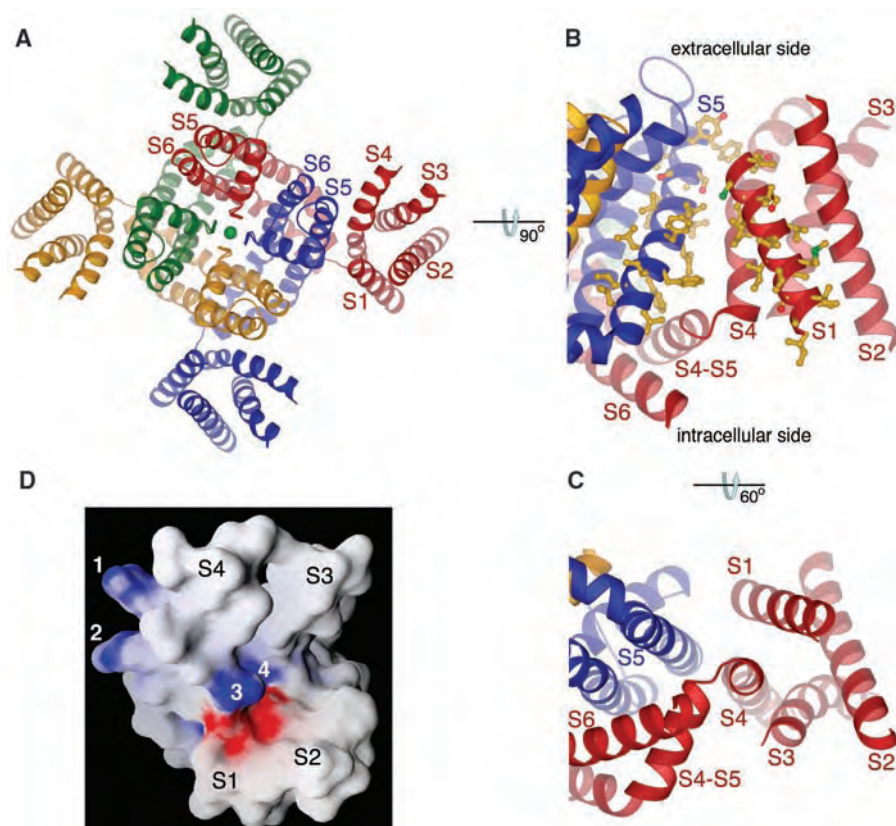


Fig. 4. Views of the integral membrane components (pore and voltage sensors) of the Kv1.2 channel. (A) Overall structure of the tetramer, viewed from the extracellular solution, shown as ribbons. Each of the four subunits is colored uniquely. The transmembrane helices S1 to S6 are labeled for the subunit colored in red. Each S4 helix (red, for example) is nearest the S5 helix of a neighboring subunit (blue, for example). (B) A close-up view of a voltage sensor and its relationship to the pore, viewed from the side. Side chains for residues on the S1 helix and the S5 helix from the neighboring subunit are shown as sticks and colored according to atom type: carbon, yellow; nitrogen, blue; oxygen, red; and sulfur, green. (C) View of the voltage sensor and pore from (B), rotated 60° around the horizontal axis to look down the S4 helix from the intracellular solution. This orientation highlights the minimal contacts between the voltage sensor and pore. (D) Surface representation of the S1-to-S4 voltage-sensor domain without the pore, viewed from the extracellular solution in the same orientation as the voltage sensor colored red in (A). The surface is colored red (negative) and blue (positive) for qualitative assessment of the electrostatic potential at the surface. The Arg residues on S4 are numbered 1 to 4. Electrostatic potential was calculated with GRASP software (43). In parts (B) and (D), the residues on S1, S2, and S4 were given complete side chains, even though some of them are modeled as polyalanine in the final coordinates. Parts (A) to (C) were generated with Molscript (42).

brane surface to extend to the phospholipid head-group layer, whereas the second is somewhat deeper. Arg residues 3 and 4 face helices S1 and S2, where they can make salt bridge interactions with acidic amino acids (Fig. 4D). These four Arg positions in the structure are in agreement with electron paramagnetic resonance (EPR) data on the KvAP channel in lipid membranes (36). Although the authors of the EPR study concluded that the Arg residues are buried, their data actually show a lipid environment for the first Arg, a lipid and water mixed environment for the second Arg, and a protein (neither lipid nor water) environment for Arg residues 3 and 4 (36, 37). The correlation between the crystal structure and EPR data argues that the first two of the four highly conserved S4 Arg residues are exposed to lipid in the open conformation of the voltage sensor (see below).

Mechanism of voltage-dependent gating. Two aspects of the Kv1.2 crystal structure suggest that we have determined an open conformation of the channel. First, the inner helix bundle (activation gate) of the pore is opened to ~ 12 Å in diameter (Fig. 3C). Second, the voltage sensors appear to be in an open position (Fig. 5, A and B); that is, when opening, the voltage sensors move the gating-charge Arg residues nearer to the extracellular side of the cell membrane. This is where we find the Arg residues in the structure if we imagine the channel embedded in a membrane—all four are above the midpoint of the membrane (Fig. 5, A and B).

How might the channel close? Mere inspection of the structure evokes a mechanism (Fig. 5A). In a closed conformation, the inner helix bundle of the pore is expected to be closed as in KcsA, and the voltage sensors are expected to be in a position that will bring the gating-charge Arg residues closer to the intracellular side of the membrane. An inward displacement of the S4 helices (downward in Fig. 5A) will bring the Arg residues toward the intracellular side of the membrane, and at the same time, it will push down on the S4-S5 linker helices. The S4-S5 linker helices will then compress the inner helices and close the pore (Fig. 3D). At a qualitative level, one can understand how a transmembrane electric field, by working on the positive Arg charges on S4, can open the pore when the membrane is positive inside (pushing the charges out) and close the pore when the membrane is negative inside (drawing the charges in).

Many important details have not been specified in the simplified description above, but the process is constrained by further data. One constraint on S4 movements in membranes comes from studies of avidin accessibility to biotin that is tethered on the KvAP channel. These studies have shown that the voltage-sensor paddle in KvAP (helices S3b and S4) is uniquely mobile and that a segment of the S4 helix moves a distance of more than

15 Å through the thickness of the membrane (23, 24). We know, for example, that positions on S4 (marked by black and blue spheres in Fig. 5B) come within a few angstroms of the extracellular and intracellular solutions, respectively, when the voltage sensors move (24). The black sphere is near the extracellular side in the open crystal structure of Kv1.2. In a closed conformation, the blue sphere would have to move to the level of the S4-S5 linker, a displacement of at least 15 Å from its position in the crystal structure. The measured accessibility of cysteine residues on S4 of the *Shaker* channel to water-soluble sulfhydryl reagents is consistent with the biotin-avidin data on KvAP (14–16). S4 movements of this magnitude would transfer the Arg residues far enough to account for the large gating charge associated with *Shaker* Kv channel opening (4) and to account for the conformational changes required to open and close the pore (Fig. 3, C and D).

Another constraint comes from the observation that an antiparallel relationship between

S3b and S4 has so far been observed in every crystal structure of KvAP (22) (see also PDB ID 2A0L), and now we observe a similar relationship between S3 and S4 in Kv1.2. We therefore suppose that S3 and S4 move together as a voltage-sensor paddle unit. We imagine that to close the channel, the paddle undergoes a motion with respect to S1 and S2, with S3 remaining “above” (on the extracellular side of S4) and S4 “below,” closer to the intracellular solution. The comparison of the voltage-sensor structure of Kv1.2 and the isolated voltage-sensor structure of KvAP (Fig. 2B, gray and blue traces) offers a suggestion of how a voltage-sensor paddle might begin to move away from its open conformation as a channel begins to close.

It has been argued that accessibility to the top (C-terminal half) of S3 from the extracellular solution in the closed conformation is inconsistent with motions of a voltage-sensor paddle (38). But these arguments are based on the perception that the top of S3 in the paddle moves near to the intracellular side and be-

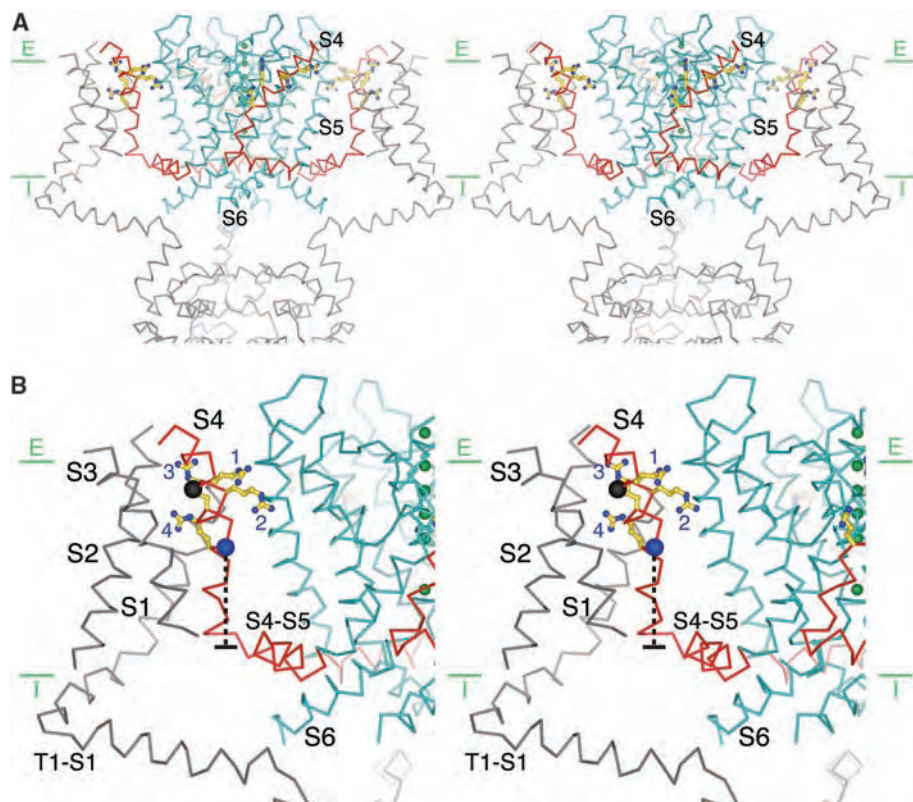


Fig. 5. Stereoview of the Kv1.2 channel showing the pore, voltage sensors, and half of the T1 domain. (A) The protein main chain is represented as a C_{α} trace. The pore is shown in cyan; the S4 helix and the S4-S5 linker in red; and voltage-sensor helices S1 to S3, the T1-S1 linker, and the T1 domain (bottom) in gray. Side chains of Arg 1 to 4 on the S4 helix are shown. α helices S1 to S3, the T1-S1 linker, and the T1 domain are removed from the subunit nearest the viewer. Green lines labeled E (extracellular) and I (intracellular) mark the approximate boundaries of a membrane 30 Å thick. (B) An enlarged stereoview of one voltage sensor is shown with the same orientation and coloring as in (A). A black sphere at position 295 highlights that the α carbon of the equivalent amino acid in KvAP approaches within a few angstroms of the extracellular solution (top) when the channel is opened at depolarized membrane voltages (positive inside), as assessed through avidin capture of tethered biotin (24). A blue sphere at position 302 shows that the equivalent position in KvAP approaches within a few angstroms of the intracellular solution (approximate distance shown as dashed line) when the channel is closed at negative membrane voltages (negative inside). The figure was generated with Molsript (42).

comes completely buried by the hydrophobic core of the lipid membrane (39, 40). In fact, the biotin-avidin studies on KvAP indicate that the top of S3 (S3b in KvAP) does not penetrate deeper than the membrane's outer leaflet (23, 24). Thus, we do not expect there to be a complete hydrophobic core covering this region in the closed conformation. The top of S3 should remain chemically and electrically near the extracellular side.

We hypothesize that S3 of the voltage-sensor paddle serves two important functions: to provide rigidity to S4 and to oppose the inward (closure) movement of the voltage sensor. The C-terminal end of S3 in the voltage-sensor paddle should be more stable at the membrane interface (than in the hydrophobic core) because it contains a mixture of hydrophobic and hydrophilic amino acids. Therefore, at negative membrane voltages, the inward movement of the paddle must always oppose the energetic preference of S3 for the interface. In this way, S3 in the voltage-sensor paddle might serve as a recoil device, causing the voltage sensor to spring to its open conformation when the membrane is depolarized.

Discussion. Crystals of the Kv1.2 K⁺ channel provide a view of a Kv channel with its voltage sensors in an apparently native conformation. The electron density for the voltage sensors is weak, but along with information from structures of KvAP, many important questions are answered by the Kv1.2 structure. In particular, three important ideas about voltage-dependent gating are conveyed.

First, a Kv channel is not composed of a rigid disk of α helices in the membrane. Rather, the voltage sensors are self-contained domains, quite independent of the pore except for their specific localized attachments (through the S4-S5 linker) that enable them to perform mechanical work on the pore. In this respect, the voltage-sensor domains are structurally analogous to the ligand-binding domains of ligand-gated ion channels, which are attached to the pore but are separate from it. Instead of being outside the membrane, as in the case of ligand-binding domains, the voltage-sensor domains are membrane-spanning. The Kv channel is the only membrane protein that we know of so far to contain separate domains within the membrane, but others will no doubt be identified in the future. A self-contained voltage sensor means that S4 is not buried in a protein-lined canaliculus. All charge shielding from the membrane and compensation by counter-charges must come from within the voltage-sensor domain itself. We see that the Arg-containing S4 helix is shielded on one face by the S1 and S2 helices of the voltage sensor and is exposed to lipid on the opposite face (Fig. 4 and Fig. 5). A self-contained voltage sensor also means that the position of the voltage sensor with respect to the pore could vary somewhat among different Kv channels. In KvAP, in

which the position of the voltage sensors is not constrained by a connection of S1 to a T1 domain, EPR studies show that S1 is mostly buried in protein, rather than exposed to lipid (36). This observation can be explained if, in KvAP, the voltage-sensor domains are repositioned slightly (i.e., rotated) to bury S1 between the voltage-sensor domain and the pore (Fig. 4A). Such a repositioning is possible without disrupting the S4-S5 linker's attachment to S6, and although the voltage-sensor paddle would be brought further out on the perimeter, the degree to which S4 is shielded would be unchanged, because the shielding is provided by S1 and S2 of the domain itself.

Second, the Kv1.2 structure shows us how conformational changes within the voltage sensors are transmitted to the pore (Fig. 3, C and D, and Fig. 5, A and B). This is an aspect of voltage-dependent gating that, until now, has eluded a mechanical explanation. The mechanism, depicted in Figs. 3 and 5, is simple: Motions of the S4 helices are transmitted to the inner helix bundle (activation gate) via the S4-S5 linker helices. This is perhaps one of the most straightforward, understandable mechanical systems observed in a protein. When inspecting the Kv1.2 structure, it is at first surprising to see that the voltage sensors are essentially "domain swapped" to the opposite side of neighboring subunits (Fig. 4A). But this arrangement actually permits the S4-S5 linker to form its mechanical attachment to the S6 inner helix, allowing the voltage sensors to perform mechanical work on the pore.

Third, in the open conformation, Arg residues 1 and 2 are on the lipid-exposed surface of the voltage sensor, and Arg residues 3 and 4 are in a position to interact with acidic amino acids inside the domain, between the voltage-sensor paddle (S3 and S4) and voltage-sensor helices S1 and S2. We think that an energetic balance between electrostatic and hydrophobic forces is important for the function of voltage sensors (25).

The Kv1.2 structure offers an explanation for many experimental results and ideas that, until now, have seemed contradictory. The original working model based on the KvAP crystal structures is substantially refined by the Kv1.2 structure, but of course this is still a working model, to be modified as new data are obtained. The question of how the voltage sensor moves from the open conformation that we now see to a closed conformation will require further study. This new structure should help in designing the next-stage experiments to test voltage-sensor movements.

References and Notes

1. F. J. Sigworth, *Q. Rev. Biophys.* **27**, 1 (1994).
2. C. M. Armstrong, F. Bezanilla, *J. Gen. Physiol.* **63**, 533 (1974).
3. F. Bezanilla, *Physiol. Rev.* **80**, 555 (2000).
4. N. E. Schoppa, K. McCormack, M. A. Tanouye, F. J. Sigworth, *Science* **255**, 1712 (1992).

5. S. K. Aggarwal, R. MacKinnon, *Neuron* **16**, 1169 (1996).
6. A. S. Seoh, D. Sigg, D. M. Papazian, F. Bezanilla, *Neuron* **16**, 1159 (1996).
7. F. Bezanilla, *J. Gen. Physiol.* **120**, 465 (2002).
8. R. Horn, *J. Gen. Physiol.* **120**, 449 (2002).
9. C. S. Gandhi, E. Y. Isacoff, *J. Gen. Physiol.* **120**, 455 (2002).
10. M. Laine, D. M. Papazian, B. Roux, *FEBS Lett.* **564**, 257 (2004).
11. F. Elinder, P. Arhem, H. P. Larsson, *Biophys. J.* **80**, 1802 (2001).
12. Y. Li-Smerin, D. H. Hackos, K. J. Swartz, *Neuron* **25**, 411 (2000).
13. N. Yang, A. L. George Jr., R. Horn, *Neuron* **16**, 113 (1996).
14. L. M. Mannuzzu, M. M. Moronne, E. Y. Isacoff, *Science* **271**, 213 (1996).
15. S. P. Yusaf, D. Wray, A. Sivaprasadarao, *Pflugers Arch.* **433**, 91 (1996).
16. O. S. Baker, H. P. Larsson, L. M. Mannuzzu, E. Y. Isacoff, *Neuron* **20**, 1283 (1998).
17. H. P. Larsson, O. S. Baker, D. S. Dhillon, E. Y. Isacoff, *Neuron* **16**, 387 (1996).
18. A. Cha, G. E. Snyder, P. R. Selvin, F. Bezanilla, *Nature* **402**, 809 (1999).
19. K. S. Glauner, L. M. Mannuzzu, C. S. Gandhi, E. Y. Isacoff, *Nature* **402**, 813 (1999).
20. D. M. Starace, F. Bezanilla, *Nature* **427**, 548 (2004).
21. Z. Lu, A. M. Klem, Y. Ramu, *Nature* **413**, 809 (2001).
22. Y. Jiang et al., *Nature* **423**, 33 (2003).
23. Y. Jiang, V. Ruta, J. Chen, A. Lee, R. MacKinnon, *Nature* **423**, 42 (2003).
24. V. Ruta, thesis, Rockefeller University (2005).
25. T. Hessa, S. H. White, G. von Heijne, *Science* **307**, 1427 (2005).
26. S. B. Long, E. B. Campbell, R. MacKinnon, *Science* **309**, 897 (2005); published online 7 July 2005 (10.1126/science.1116269).
27. Y. Liu, M. Holmgren, M. E. Jurman, G. Yellen, *Neuron* **19**, 175 (1997).
28. N. E. Schoppa, F. J. Sigworth, *J. Gen. Physiol.* **111**, 295 (1998).
29. D. H. Hackos, T. H. Chang, K. J. Swartz, *J. Gen. Physiol.* **119**, 521 (2002).
30. O. Yifrach, R. MacKinnon, *Cell* **111**, 231 (2002).
31. M. Sukhareva, D. H. Hackos, K. J. Swartz, *J. Gen. Physiol.* **122**, 541 (2003).
32. Z. Lu, A. M. Klem, Y. Ramu, *J. Gen. Physiol.* **120**, 663 (2002).
33. M. Laine et al., *Neuron* **39**, 467 (2003).
34. D. J. Elliott et al., *EMBO J.* **23**, 4717 (2004).
35. Y. Murata, H. Iwasaki, M. Sasaki, K. Inaba, Y. Okamura, *Nature* **435**, 1239 (2005).
36. L. G. Cuello, D. M. Cortes, E. Perozo, *Science* **306**, 491 (2004).
37. R. MacKinnon, *Science* **306**, 1304 (2004).
38. C. S. Gandhi, E. Clark, E. Loots, A. Pralle, E. Y. Isacoff, *Neuron* **40**, 515 (2003).
39. H. C. Lee, J. M. Wang, K. J. Swartz, *Neuron* **40**, 527 (2003).
40. C. Gonzalez, F. J. Morera, E. Rosenmann, O. Alvarez, R. Latorre, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5020 (2005).
41. A. T. Brunger et al., *Acta Crystallogr.* **D54**, 905 (1998).
42. P. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).
43. A. Nicholls, K. A. Sharp, B. Honig, *Proteins* **11**, 281 (1991).
44. We thank A. Lee, V. Ruta, and members of the MacKinnon laboratory for helpful discussions; R. Jain for initial experiments with lipids; R. Dutzler for assistance with data collection; O. Pongs for Kv1.2 DNA; J. Trimmer for $\beta 2$ subunit DNA; Brookhaven National Laboratory (National Synchrotron Light Source beamlines X25 and X29) and the Swiss Light Source (beamline PX1) staff for assistance in data collection; and W. Chin for help with manuscript preparation. This work was supported in part by NIH grant no. GM43949 to R.M. R.M. is an Investigator in the Howard Hughes Medical Institute. Atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession ID 2A79.

17 June 2005; accepted 5 July 2005

Published online 7 July 2005;

10.1126/science.1116270

Include this information when citing this paper.

Triangular and Fibonacci Number Patterns Driven by Stress on Core/Shell Microstructures

Chaorong Li, Xiaona Zhang, Zexian Cao*

Fibonacci number patterns and triangular patterns with intrinsic defects occur frequently on nonplanar surfaces in nature, particularly in plants. By controlling the geometry and the stress upon cooling, these patterns can be reproduced on the surface of microstructures about 10 micrometers in diameter. Spherules of the Ag core/SiO_x shell structure, possessing markedly uniform size and shape, self-assembled into the Fibonacci number patterns (5 by 8 and 13 by 21) or the triangular pattern, depending on the geometry of the primary supporting surface. Under proper geometrical constraints, the patterns developed through self-assembly in order to minimize the total strain energy. This demonstrates that highly ordered microstructures can be prepared simultaneously across large areas by stress engineering.

Growing highly ordered micro- and nano-structures in a designed pattern with acceptable size and shape uniformity presents an enormous challenge to materials scientists. Self-assembly offers a promising approach, in that it allows the production of ordered structures without the need for high-precision patterning equipment (1, 2). Various routes that can lead to self-assembly have been used, and colloidal chemistry in particular has produced some impressive results (2–5). Nevertheless, applying physical principles at a macroscopic scale rather than relying on intermolecular forces would be more practicable for microfabrication, especially when control of purity and uniformity is of concern. Recently, stress engineering has been used to make ordered structures on large scales (6–10). Typically, this has been undertaken on planar supporting surfaces. The possibility of exploiting geometrical constraints has not been widely realized.

The geometry of receptacles should conceivably play an important, if not decisive, role in determining the arrangement of the growing, repeated botanic elements such as florets, sepals, and seeds, which are positioned in fascinating, highly ordered patterns (11, 12). Triangular patterning, also referred to as hexagonal close-packing, is observed on large, flat surfaces. For example, the arrangement of thorns in the central part of the cactus *Opuntia catingicola* is a nearly perfect triangular lattice. On a spherical receptacle, the triangular pattern has to adopt pentagonal or heptag-

onal defects to meet the conditions imposed by Euler's rule for convex polyhedrons (13). It is also noteworthy that spherical receptacles are the preferred support for botanic elements that grow in tandem, such as the florets on a dandelion. The Fibonacci number pattern, in which the basic units can be grouped into clockwise and counterclockwise spirals specified by two neighboring numbers in the Fibonacci sequence, is manifested by many plants, where the receptacle is generally a disk or conical surface, and the growing elements develop consecutively from primordia. The most familiar examples include the sunflower, daisy, and pineapple. In a sunflower, the Fibonacci number pattern can range from 21 by 34 to 89 by 144 in a giant flower head. Of course, in all the natural patterns, deviations from such patterns often occur owing to the irregularity to the receptacles, phototropism, and other conditional factors.

Both triangular and Fibonacci number patterns were produced by stress-driven self-assembly on the surface of Ag core/SiO_x shell microstructures, where the value of x is only slightly smaller than 2.0 as confirmed by energy-dispersive x-ray spectroscopy; the shell structures were prepared primarily at a high temperature and then allowed to cool down. The Ag core/SiO_x shell structures of a few micrometers in dimension were grown on polycrystalline sapphire substrate by coevaporating a mixed powder of Ag₂O and SiO in a mass ratio of 4:1 (fig. S1). The substrate temperature was maintained at 1270 K, a temperature deliberately chosen to be slightly above the melting point of silver (1234.8 K) but far below that of SiO₂ (1883 K). The Ag core/SiO_x shell structures formed at this stage through high-temperature interdiffusion due to

the poor miscibility of Ag with silicon oxide. Depending on the local wetting condition, a liquid drop of Ag enveloped by a thin layer of SiO_x can adopt, loosely speaking, both nearly spherical or a flattened, conical morphology. Upon cooling (fig. S2), large compressive stresses develop in the shell layer due to the sizable mismatch of expansion coefficients between the core and the shell materials. In comparison to silicon oxide, Ag in the melt is a more compliant material. For silver, the coefficient of thermal expansion (at 1100 K) is $\alpha_{\text{Ag}} = 27.1 \times 10^{-6}$, whereas for SiO₂, $\alpha_{\text{SiO}_2} = 0.45 \times 10^{-6}$. For a temperature drop $\Delta T \approx 800$ K, the equibiaxial stress in the prebuckling state is ~ 2.0 GPa, as estimated from the relation $\sigma = E\Delta\alpha\Delta T/(1 - \nu)$, where $\Delta\alpha = \alpha_{\text{Ag}} - \alpha_{\text{SiO}_2}$, E is the Young's modulus, and ν is the Poisson ratio (7). For SiO₂, $E = 75$ GPa, and $\nu = 0.17$ (14). For a typical SiO₂ shell with a thickness of 150 nm, the strain energy is ~ 6.2 J/m² as calculated from the formula $U = [E/(1 - \nu)]\Delta\alpha^2\Delta T^2t$, where t is the shell thickness. The strain energy is much larger than the surface energy, which is also an important factor in modifying the morphology of the silicon oxide surface at the given temperatures, and hence is the dominant driving force for the formation of patterns discussed below.

To minimize the temperature inhomogeneity in the prebuckled state, the cooling rate must be carefully controlled. A large cooling rate will cause the silver core to contract too fast, producing shriveled or even broken shells, as seen under the scanning electron microscope (SEM) (fig. S3). With a moderate cooling rate (fig. S2) that allows the temperature to be reduced to 750 K in 4 min or more, large stress accumulates in the SiO_x shell, yet the shell remains attached to the silver core and, importantly, has a continuous, convex surface. This heavily stressed layer is unstable against the formation of strained domains that reduce the strain energy (15). It will also modify the nucleation barrier and diffusivity for the subsequent growth at low supersaturation—the evaporation continues but at lowered temperatures, and consequently at a reduced vapor pressure. The resulting distribution of three-dimensional (3D) condensates, which are spherules that also possess a Ag core/SiO_x shell structure, demonstrates the pattern of stress. Consequently, we observed a tessellation of the surface of the primary Ag core/SiO_x shell structure by patterns in which the spherules appear as vortices. Figure 1, A to C, shows SEM images of primary Ag core/SiO_x shell structures, typically ~ 10 μm in diameter and with a shell thickness of ~ 150 nm. The shell structures all have a nearly spherical surface, on which a triangular pattern of spherules of markedly uniform size and shape

Institute of Physics, Chinese Academy of Sciences, Post Office Box 603, 100080 Beijing, China.

*To whom correspondence should be addressed. E-mail: zxcao@aphy.iphy.ac.cn

formed. This decoration of the most stressed sites in the close-packing pattern, with the spherules, indicates the tendency of the primary surface to minimize its strain energy. The finite, nonplanar geometry is expected to impose a greater constraint on the stress pattern. Because congruent triangles cannot tile a spherical surface, the triangular pattern inevitably includes some defects—five- or sevenfold triangles around a vortex. Natural occurrences of this self-assembled pattern are ubiquitous on flower heads that are nearly spherical in shape (Fig. 1D).

On a given spherical surface, the triangular pattern can be characterized solely by the parameter δ , the arc separation between the nearest vortices, provided that the defects occur only sporadically. Roughly speaking, the arc separation δ corresponds to the wavelength of the buckling mode in the planar geometry, which scales approximately as the square root of the stiffness of the buckling material (7, 16). For a spherical shell, the radius R enters into the expression of the total strain energy as R^2 ,

and the number of stressed vortices is $\propto R^2/\delta^2$; therefore, the arc separation δ in the triangular pattern will show only a weak dependence on the size of the primary spherical surface. In contrast, the stiffness D of the shell is related to the thickness t through $D = Et^3/[12(1 - \nu^2)]$. Hence, the thicker the shell layer, the larger the arc separation δ for the pattern. As expected, the densest pattern is observed in the shell structure shown in Fig. 1A, which has the best image quality because the SiO_x layer is very thin. In Fig. 1, B and C, the SiO_x shell layer is thicker, as evidenced by the deteriorated image quality—the poorer conductivity of the thicker SiO_x layer caused a severe charging of the sample. The primary core/shell in Fig. 1C also has a very large diameter, and therefore the triangular pattern on it has the largest arc separation, $\delta \approx 1.78 \mu\text{m}$.

When the primary Ag core/ SiO_x shell structure better wetted the substrate, it adopted a flattened, conical shape, and the spherules arranged themselves in distinct spirals. On a

core/shell of $\sim 9.5 \mu\text{m}$ in lateral dimension (Fig. 2, A and B), the 92 spherules can be assigned to eight clockwise and five counterclockwise spirals, and thus they form a 5 by 8 Fibonacci number pattern. In Fig. 2C, the lateral dimension for the core/shell is about $18 \mu\text{m}$, and the spherules, about 230 in total, form a 13 by 21 Fibonacci number pattern, similar to the one on the cactus *Mammillaria nejapensis* (Fig. 2D).

The formation of patterns of stressed sites on a surface to reduce the total strain energy can be modeled by the general minimal Riesz energy point configuration problem, for which only numerical solutions are available (17, 18). For a spherical supporting surface where the point number is large enough but less than 500, numerical simulation shows that a triangular pattern with 12 disclinations (fivefold defects) is always produced (18). From Fig. 1, however, we see that sevenfold defects occur frequently in a real system, reminding us of the imperfection of the supporting surfaces either grown in nature or prepared in the laboratory.

We have demonstrated that highly ordered patterns can be fabricated on core/shell structures through stress-driven self-assembly induced by controlled cooling. The key to obtaining a particular pattern is to control both the geometry of the primary core/shell and the thickness of the more rigid shell layer. The final configuration of the pattern can be explained by the principle of minimum strain energy under given geometrical constraints. Moreover, it demonstrates a fabrication technique for the mass production of mesoscale structures on a large area and a way to incorporate desirable “defects” simultaneously. Our results indicate a way to control the pattern assembly process and raise the possibility of designing an entire family of patterns by stress engineering.

References and Notes

1. N. Bowden, S. Brittain, A. G. Evans, J. W. Hutchinson, G. M. Whitesides, *Nature* **393**, 146 (1998).
2. Y. A. Vlasov, X. Z. Bo, J. C. Sturm, D. J. Norris, *Nature* **414**, 289 (2001).
3. S. Aggarwal et al., *Science* **287**, 2235 (2000).
4. J. P. Hill et al., *Science* **304**, 1481 (2004).
5. P. Y. Su, J. C. Hu, S. L. Cheng, L. J. Chen, J. M. Liang, *Appl. Phys. Lett.* **84**, 3480 (2004).
6. R. Adelung et al., *Nat. Mater.* **3**, 375 (2004).
7. X. Chen, J. W. Hutchinson, *Scr. Mater.* **50**, 797 (2004).
8. Y. D. Yin et al., *Science* **304**, 711 (2004).
9. X. N. Zhang, C. R. Li, Z. Zhang, *Mater. Lett.* **58**, 1917 (2004).
10. H. Lee, J. A. Johnson, M. Y. He, J. S. Speck, P. M. Petroff, *Appl. Phys. Lett.* **78**, 105 (2001).
11. I. Adler, D. Barabe, R. V. Jean, *Ann. Bot. (London)* **80**, 231 (1997).
12. H. Weyl, *Symmetry* (Princeton Univ. Press, Princeton, NJ, 1952).
13. P. R. Cromwell, *Polyhedra* (Cambridge Univ. Press, Cambridge, 1997).
14. G. W. C. Kaye, T. H. Laby, *Tables of Physical and Chemical Constants* (Longman Science and Technology, Sussex, UK, 1995).
15. M. Hanbücken, J. Deville, *Stress and Strain in Epitaxy: Theoretical Concepts, Measurements and Applications* (Elsevier, Amsterdam, 2001).
16. S. P. Timoshenko, J. M. Gere, *Theory of Elastic Stability* (McGraw-Hill, New York, ed. 2, 1961).
17. E. R. Altschuler et al., *Phys. Rev. Lett.* **78**, 2681 (1997).

Fig. 1. Self-assembled triangular pattern on nearly spherical receptacles. (A) The diameter of the primary Ag core/ SiO_x shell is $D \approx 9.1 \mu\text{m}$, and the arc separation between the spherules is $\delta \approx 0.82 \mu\text{m}$; (B) $D \approx 6.6 \mu\text{m}$, $\delta \approx 1.0 \mu\text{m}$; and (C) $D \approx 11.5 \mu\text{m}$, $\delta \approx 1.78 \mu\text{m}$. (D) The spherules shown make a perfect tessellation of the surface with equilateral triangular facets. (E) A bud of *Succisa pratensis* Moench with close-packing florets is presented for comparison. Intrinsic defects are highlighted.

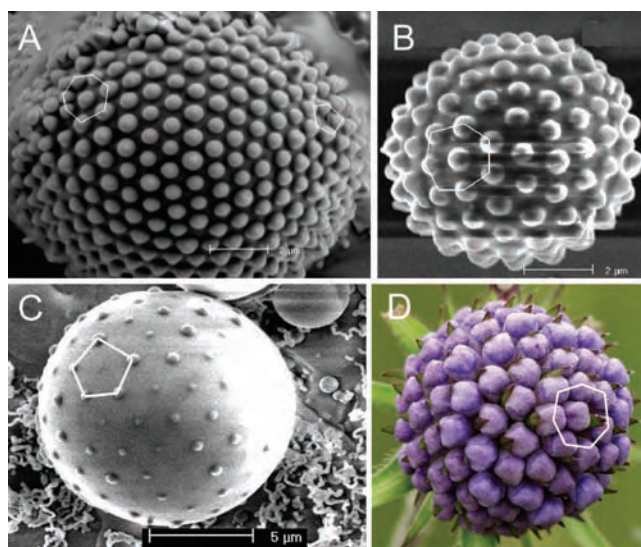
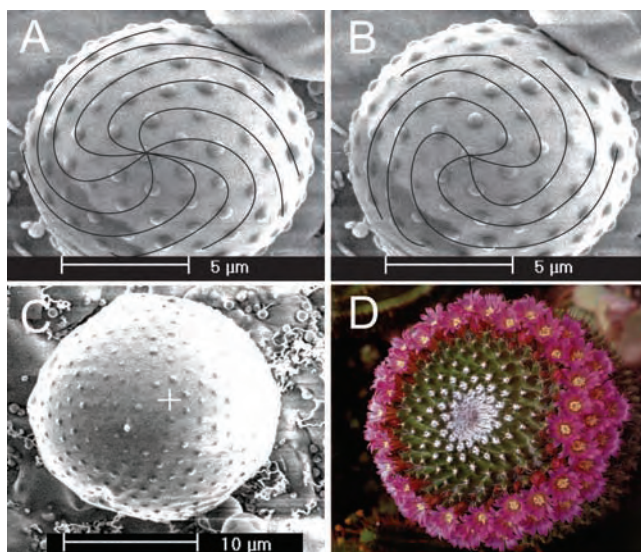


Fig. 2. Formation of Fibonacci number patterns on conical receptacles. The self-assembled spherules on a conical Ag core/ SiO_x shell of $\sim 9.5 \mu\text{m}$ in lateral dimension can be grouped into eight clockwise (A) and five counterclockwise spirals (B). (C) Taking the marked point (+) as the ending of the spirals, the ~ 230 spherules on a $\sim 17.8 \mu\text{m}$ -large conical surface can be arranged into 13 counterclockwise and 21 clockwise spirals, similar to the arrangement of spherules in the flowering cactus *Mammillaria nejapensis* (D).



18. M. Bowick, A. Cacciuto, D. R. Nelson, A. Travesset, *Phys. Rev. Lett.* **89**, 185502 (2002).
19. We acknowledge support from the China State Key Projects of Basic Research (grant 2002CB613500), the National High Technology Research and Development Program of China (grant 2003AA302170),

and the National Natural Science Foundation of China (grants 60306009, 50472070, and 50272081).

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5736/909/DC1

Materials and Methods
Figs. S1 to S3

11 April 2005; accepted 23 June 2005
10.1126/science.1113412

A Crossover in the Mechanical Response of Nanocrystalline Ceramics

Izabela Szlufarska,^{1*} Aiichiro Nakano,² Priya Vashishta²

Multimillion-atom molecular dynamics simulation of indentation of nanocrystalline silicon carbide reveals unusual deformation mechanisms in brittle nanophase materials, resulting from the coexistence of brittle grains and soft amorphous grain boundary phases. Simulations predict a crossover from intergranular continuous deformation to intragrain discrete deformation at a critical indentation depth. The crossover arises from the interplay between cooperative grain sliding, grain rotations, and intergranular dislocation formation similar to stick-slip behavior. The crossover is also manifested in switching from deformation dominated by indentation-induced crystallization to deformation dominated by disordering, leading to amorphization. This interplay between deformation mechanisms is critical for the design of ceramics with superior mechanical properties.

The great interest in nanostructured ceramics originates from the observations and expectations of unique mechanical properties (1–3) in these materials. Examples include very high hardness, high fracture toughness, and superplastic behavior in normally brittle ceramics. Silicon carbide is of particular interest because of its potential technological applications in high-temperature structural and electronic components (4). Although enhanced mechanical properties are often associated with the reduction in grain sizes, it has recently been conjectured (5) that nanostructured ceramics might exhibit an inverse Hall-Petch effect; that is, hardness decreases when grain size decreases in the nanoscale grain-size regime. Such peculiar behavior has been observed in ductile nanophase materials (e.g., nanostructured metals) with porous grain boundaries (GBs) by means of simulations (6–8) and experiments (9). The behavior was attributed to a crossover from dislocation-mediated plasticity for large grain size to GB sliding for small grain size (10, 11). A similar mechanistic understanding in ceramics is still lacking.

In contrast with nanostructured metals, nanostructured ceramics have an increased volume fraction of disordered intergranular films, which

are observed both experimentally (12) and by means of molecular dynamics (MD) simulations (13, 14). In particular, for brittle ceramics such as SiC, mechanical properties such as toughness are essentially determined by soft (often amorphous) GB phases (12). Recent experiments (15) of nanoindentation of nanocrystalline SiC (n-SiC) films with grain sizes of 5 to 20 nm have shown “superhardness,” i.e., hardness largely exceeding that of a bulk crystalline SiC (3C-SiC). The experimental hardness was shown to be sensitive to the grain size and the fraction of the amorphous GB phase. However, their effects on mechanical responses at the atomistic level are largely unknown.

The MD simulations consisted of a $625 \times 625 \times 535$ Å³ n-SiC substrate containing

18.7 million atoms, which had randomly oriented grains with diameters averaging 8 nm and a density of 2.97 g/cm³ at a temperature of 300 K (16). Structural ordering in GBs is analyzed by means of a partial pair distribution function $g(r)$, which quantifies the probability of finding two atoms at an interatomic distance r . The function $g(r)$ for Si-C pairs is plotted in Fig. 1A (solid line), and it reveals a lack of long-range order, similar to that of bulk amorphous SiC (a-SiC) (dashed line). This is in contrast to a sharp-peak structure of $g(r)$ in a 3C-SiC shown in the inset of Fig. 1A. Visual inspection of the substrate reveals the presence of highly disordered GBs with a more or less uniform thickness. The n-SiC can be thought of as a substrate with two coexisting phases: crystalline inside the grains and amorphous in the GBs. Amorphous GBs were also observed in experimentally sintered n-SiC (17, 18).

To shed light on the atomistic mechanisms underlying mechanical response of n-SiC, we indented the substrate with a square-base indenter of size $160 \times 160 \times 72$ Å³. Nanoindentation is a unique local probe to measure mechanical properties of materials (19–22). Even though experimental indenters are round on this scale, a square-base indenter helps to maximize the applied stress and the localized plastic flow in the material (23) on length and time scales available to simulations. The resulting load-displacement (P - h) curve is shown in Fig. 1B (solid line), together with two unloading curves (dashed lines). The P - h response exhibits four characteristic regimes.

Regime 1 is entirely elastic and ends at $h = 7.5$ Å. Regime 2 extends up to the crossover depth $h_{CR} \sim 14.5$ Å and is characterized by a very small hysteresis during unloading as com-

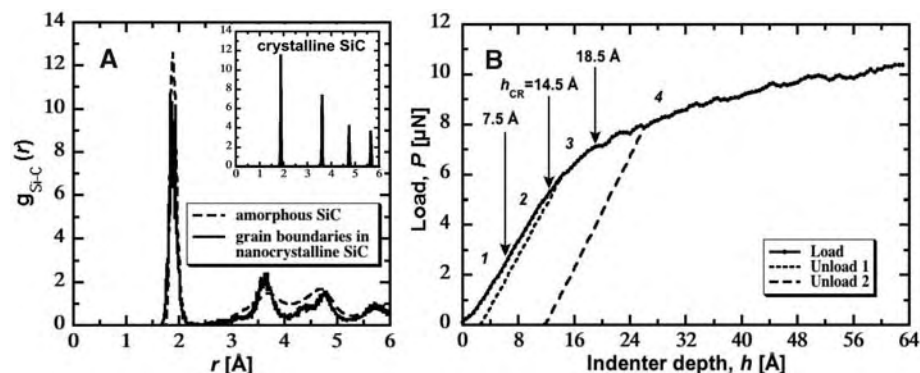


Fig. 1. (A) Si-C partial pair distribution function $g_{Si-C}(r)$ of GBs in n-SiC (solid line) resembles that of a-SiC (dashed line) in contrast to 3C-SiC (inset). (B) Load-displacement curve of 18.7-million-atom n-SiC, where numbers 1 to 4 mark different regimes in the response. Regimes 1 and 2 are characterized by continuous cooperative grain motion; in regime 3 grains become decoupled from one another; in regime 4 discrete intergrain response is turned on, which is manifested in the small load drops in the P - h curve.

¹Department of Materials Science and Engineering, 1509 University Avenue, University of Wisconsin, Madison, WI 53706–1509, USA. ²Department of Chemical Engineering and Materials Science, 3651 Watts Way, University of Southern California, Los Angeles, CA 90089–0242, USA.

*To whom correspondence should be addressed. E-mail: izabela@engr.wisc.edu

pared with a much more pronounced plastic yield at h_{CR} . The small plastic flow is related to the nonequilibrium structure of the amorphous interfaces, which can relax within a short migration distance (24). A similar effect has been observed during MD simulations of bulk a-SiC (25). Because up to h_{CR} the amorphous “cementlike” GBs hold the grains together, regimes 1 and 2 are characterized by cooperative continuous intergranular response. The cooperative motion of grains is identified by analyzing atomic displacements. It involves both formation of mesoscopic shear planes involving several grains and coupling to grains outside of the area directly beneath the indenter to form an extended elastic zone. Formation of mesoscopic shear planes has been previously observed in MD simulations of nanocrystalline Ni (26).

Regime 3 starts when amorphous GBs yield plastically at $h_{CR} \sim 14.5$ Å and henceforth grains are effectively decoupled from one another. The substrate contains a small number of nanopores, which collapse under the indenter, thereby reducing the yield stress. In experimental nanoceramics, the volume fraction of pores can be as high as 20% (27, 28). The crystalline phase within the grains does not yield until the onset of regime 4 at $h = 18.5$ Å. This response is distinct from that of nanostructured metals, in which a dislocation within the grain is nucleated at the onset of substrate yielding (29–33). Discrete plastic events, such as a dislocation glide, take place within the grains in close proximity to the indenter and are reflected in the rougher character of the P - h curve. Similar periodic load drops have been observed for the nanoindentation in bulk 3C-SiC. In the case of n-SiC, the load drops are much less pronounced than in 3C-SiC (25), because the calculated load is averaged over a few grains covered by the indenter and the discrete events in a grain are decoupled from those in the neighboring grains. The decoupling from grains not lying directly beneath the indenter is shown in Fig. 2A, in which atoms are color coded by ΔR_{CM} , the total displacement of the centers of mass of individual grains from their positions before indentation. The localization of deformation is much more pronounced in regimes 3 and 4 than in 1 and 2.

The crossover from cooperative mechanical response of coupled grains to discrete intergranular events is shown in Fig. 2B, where we plot the average displacement $\langle \Delta r_{CM} \rangle$ of the grains' centers of mass calculated relative to the previous indentation step. The coupling of grains in regime 1 is manifested as the peak of $\langle \Delta r_{CM} \rangle$ at $h \sim 7.5$ Å; in regime 2, the coupling slowly decreases until it becomes negligible at the onset of regime 3. In regime 4 ($h \geq 18.5$ Å), the motion of grains is decoupled and $\langle \Delta r_{CM} \rangle$ oscillates around a constant value of ~ 0.15 Å. The oscillations in $\langle \Delta r_{CM} \rangle$ are correlated with the small load drops in the P - h response. They correspond to the

interplay between the sliding of individual grains, grain rotations, dislocation formation inside the grains, and coupling of grains in the extended elastic zone underneath the plastic one. Examples of discrete plastic events are shown in Fig. 2, C to E. In Fig. 2C, we visualize displacements of atoms at $h = 29$ Å relative to their positions at the previous indentation step. Figure 2, D and E, for $h = 26.5$ and 27 Å, respectively, show a unit dislocation with Burgers vector $\frac{1}{2}[101]$ gliding in $(\bar{1}11)$ plane through a grain structure containing Shockley partial dislocations and stacking faults (SFs). These SFs exist on every second $\{111\}$ plane below the green line in Fig. 2, D and E, forming a thin bilayer of hexagonal (wurtzite) structure in the face-centered cubic (zinc-blende) structure of crystalline grains. This type of defect is similar to twin boundaries observed in indentation experiments of nanocrystalline aluminum (34), where SFs are formed on the neighboring $\{111\}$ planes. Shockley partial dislocations, SFs, and twins have also been observed experimentally in 3C-SiC (35, 36).

The crossover and localization of deformation are also manifested in the rotation of grains.

Figure 3, A and B, show atomic arrangements in regimes 1 and 4, respectively. To highlight the rotational motions of selected grains, we mark as blue a rectangular box inside the n th grain and calculate its rotation angle $\alpha(n)$ averaged over all atoms inside the box. The standard deviation $\Sigma(n)$ from the mean $\alpha(n)$ is a measure of the grain's deformation. After plastic yielding of the amorphous GBs at h_{CR} , grain 52, lying directly beneath the indenter, experiences a tremendous shear and undergoes both large rotation and deformation (black curves in Fig. 3C). On the other hand, a neighboring grain 71, not lying directly under the indenter, does not rotate or deform during indentation (blue curves in Fig. 3C). This localization of deformation is due to decoupling of grains in regimes 3 and 4. Some of the grains in the indenter vicinity exhibit a behavior similar to stick-slip behavior. For example, grain 36 (red curves in Fig. 3C) shows a sudden jump in rotation $\alpha(36)$ and deformation $\Sigma(36)$ at $h \sim 12$ Å, which is accompanied by a release of accumulated shear stress through a dislocation glide. After the stress has been released, the grain mainly rotates [increasing $\alpha(36)$ and nearly constant $\Sigma(36)$].

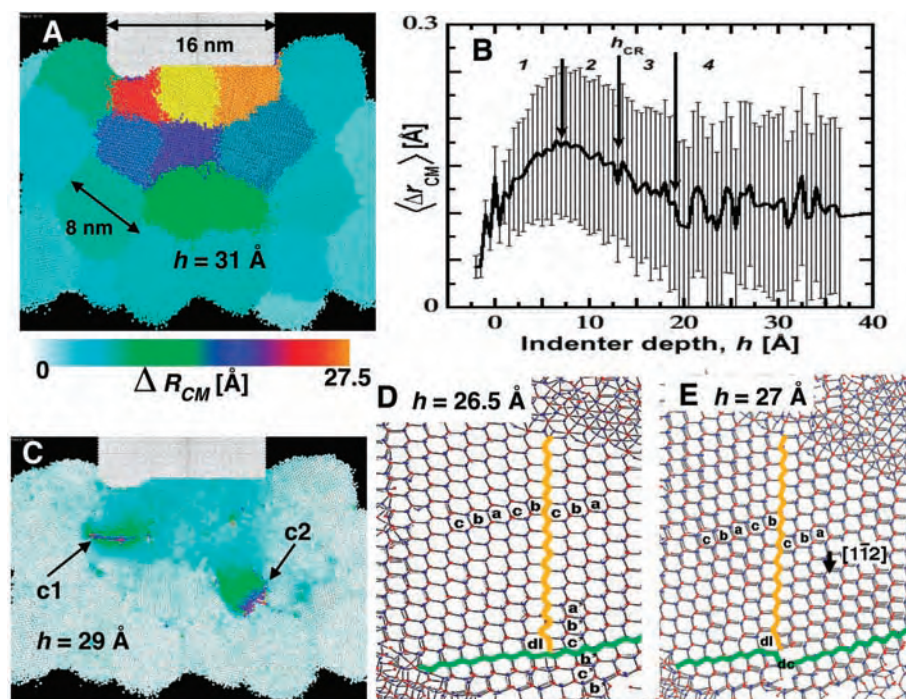


Fig. 2. (A) Grains near the indenter at $h = 31$ Å (regime 4). Atoms are color coded by ΔR_{CM} , the total displacement of grain's center of mass (CM), relative to the initial system. Decoupling of grains results in a pronounced localization of damage beneath the indenter. (B) Mean displacement $\langle \Delta r_{CM} \rangle$ of grains' CMs at every indentation step and its standard deviation. (C) Load drop at $h = 29$ Å is correlated with dislocations gliding inside the grains (c1) and at GBs (c2). Atoms are color coded by their displacement from the previous indentation position. (D and E) The (110) projection of a grain's structure before (D) and after (E) a glide of a unit dislocation with Burgers vector $\frac{1}{2}[101]$ along SF in the $(\bar{1}11)$ plane. The initial structure (D) shows a SF, abcbca, along the orange line. The SF is terminated with a Shockley partial dislocation line (dl) along the $[110]$ direction at the boundary (green line) between cubic (zinc-blende) and hexagonal (wurtzite) structure, the latter denoted with the stacking b'c'b'. In (E), a black arrow in the $[112]$ direction is a projection of the Burgers vector of the perfect dislocation. After the glide, the dislocation core (dc) becomes pinned at the zinc-blende/wurtzite boundary.

The observed crossover is correlated to a competition between crystallization and disordering in the substrate. To study this effect, we analyzed topological disorders based on the analysis of rings, where a ring is defined as the shortest closed path of alternating Si-C atomic bonds (37). Each atom in a perfect 3C-SiC has 12 unique threefold rings (these are ordered atoms), and a topological disorder is reflected in the presence of rings that are not threefold (disordered atoms). In Fig. 4A, all atoms that remain ordered from the beginning up to $h = 18.5$ Å

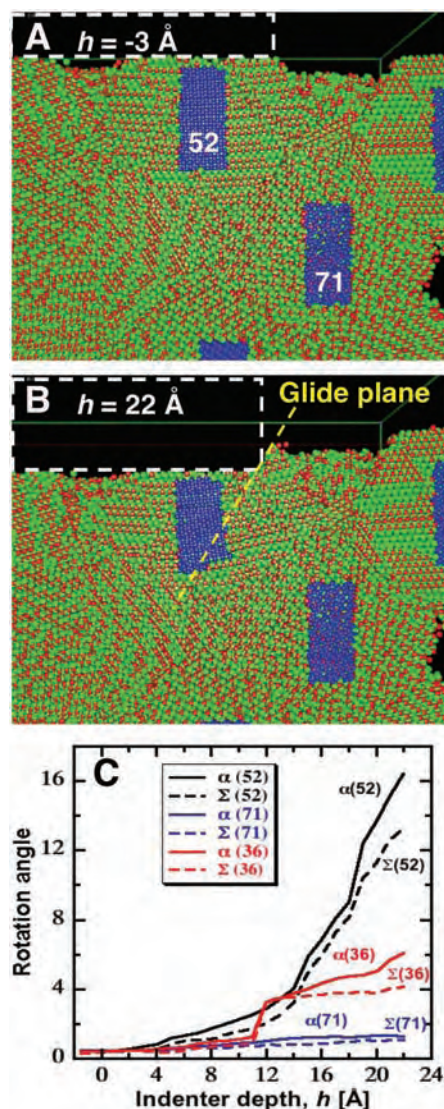


Fig. 3. Rotation of grains. (A and B) Si (green) and C (red) atoms shown in regimes 1 (A) and 2 (B) (white dashed lines depict the position of the indenter). Grain 52 experiences enormous shear beyond h_{CR} and it undergoes large rotation and deformation. (C) Rotation angle $\alpha(n)$ and its standard deviation $\Sigma(n)$ of selected grains, $n = 52, 71$, and 36. Grain 52 undergoes large rotation and deformation (black curve). Localization of deformation is reflected in the lack of rotation of grain 71 (blue curve). Grain 36 (red curve) exhibits behavior similar to stick-slip behavior at h_{CR} .

are white, and all atoms that remain disordered in the same range are yellow. The other atoms (blue and red) undergo changes in topological order, where the two competing deformation mechanisms are: ordering (crystallization) of atoms at the edges of GBs (blue) and disordering of atoms (red) in the grains. Crystallization of n-SiC has been observed experimentally during densification in the sintering process (38), and it may be responsible for the interfacial migration and small plastic flow in regime 2 of the P - h response. Even though both mechanisms are operating in the entire range of h , there is a clear switch from deformation dominated by crystallization in regimes 1 and 2 to deformation dominated by disordering in regimes 3 and 4 (Fig. 4B), in which the fraction of disordered atoms takes the minimum value at h_{CR} . The discussed disordering of atoms has been previously shown to lead to solid-state amorphization in a bulk 3C-SiC (23).

Our estimate of n-SiC hardness (defined as maximum load divided by the cross-sectional area of the indenter) of 39 GPa is in agreement

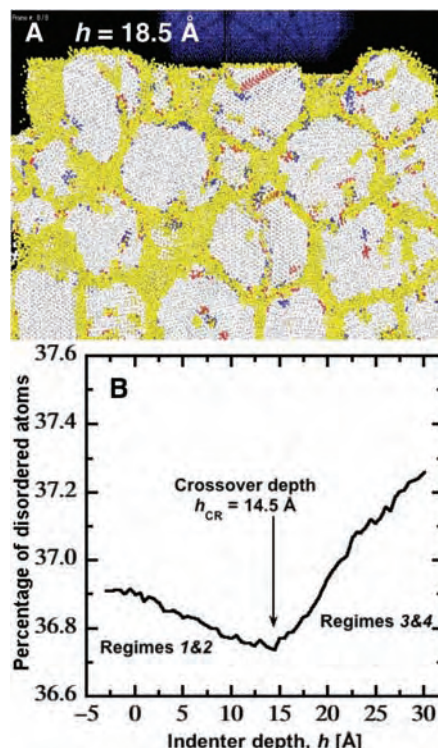


Fig. 4. Competition between crystallization and disordering in n-SiC quantified by means of shortest path ring analysis. (A) White, atoms ordered in the entire range of h up to 18.5 Å; yellow, atoms disordered in the entire range; blue, atoms changed from disordered to ordered (crystallized); red, atoms changed from ordered to disordered. (B) Percentage of disordered atoms as a function of indentation depth. The crossover at h_{CR} is correlated with the switch from dominating crystallization in regimes 1 and 2 to dominating disordering in regimes 3 and 4.

with experimental value of “superhardness” of 30 to 50 GPa for grain sizes of 5 to 20 nm (15). These experimental measurements are sensitive to the grain size and to the fraction of amorphous intergranular phase (15, 39, 40). This indicates an essential role of the interplay between discrete (crystalline) intragranular and continuous (amorphous) intergranular responses as a function of the length scale in determining mechanical properties, in particular for promising potential applications in advanced superhard nanostructured coatings, high-speed machining and tooling, or as potential materials for bio-implants. The two-phase character of nanostructured ceramics results in a crossover between the two aforementioned responses, and this crossover sheds light on the competition between different deformation mechanisms underlying design and fabrication of nanostructured ceramics with enhanced mechanical properties.

References and Notes

1. Y. Zhao *et al.*, *Appl. Phys. Lett.* **84**, 1356 (2004).
2. S. Zhang, D. Sun, Y. Fu, H. Du, *Surf. Coat. Technol.* **167**, 113 (2003).
3. R. W. Siegel, in *Nanomaterials: Synthesis, Properties and Applications*, A. S. Edelstein, R. C. Cammarata, Eds. (IOP, Bristol, 1996), pp. 201–218.
4. R. Madar, *Nature* **430**, 974 (2004).
5. J. Li, S. Yip, *Comput. Model. Eng. Sci.* **3**, 219 (2002).
6. J. Schiotz, F. D. D. Tolla, K. W. Jacobsen, *Nature* **391**, 561 (1998).
7. H. V. Swygenhoven, *Science* **296**, 66 (2002).
8. S. Yip, *Nature* **391**, 532 (1998).
9. J. A. Knapp, D. M. Follstaedt, *J. Mater. Res.* **19**, 218 (2004).
10. J. Schiotz, T. Vegge, F. D. D. Tolla, K. W. Jacobsen, *Phys. Rev. B* **60**, 11971 (1999).
11. Z. Shan *et al.*, *Science* **305**, 654 (2004).
12. D. Chen, X. F. Zhang, R. O. Ritchie, *J. Am. Ceram. Soc.* **83**, 2079 (2000).
13. P. Keblinski, S. R. Phillpot, D. Wolf, H. Gleiter, *Acta Mater.* **45**, 987 (1997).
14. P. Keblinski, S. R. Phillpot, D. Wolf, H. Gleiter, *Phys. Rev. Lett.* **77**, 2965 (1996).
15. F. Liao, S. L. Girshick, W. M. Mook, W. W. Gerberich, M. R. Zachariah, *Appl. Phys. Lett.* **86**, 171913 (2005).
16. Materials and methods are available as supporting material on Science Online.
17. D. Chen, M. E. Sixta, X. F. Zhang, L. C. D. Johnghie, R. O. Ritchie, *Acta Mater.* **48**, 4599 (2000).
18. S. Guicciardi, D. Sciti, C. Melandri, A. Bellosi, *J. Am. Ceram. Soc.* **87**, 2101 (2004).
19. W. C. Oliver, G. M. Pharr, *J. Mater. Res.* **7**, 1564 (1992).
20. J. B. Pethica, R. Hutchings, W. C. Oliver, *Philos. Mag.* **A 48**, 593 (1983).
21. W. D. Nix, *Met. Trans. A* **20A**, 2217 (1989).
22. D. Feichtinger, P. M. Derlet, H. V. Swygenhoven, *Phys. Rev. B* **67**, 024113 (2003).
23. I. Szlufarska, R. K. Kalia, A. Nakano, P. Vashishta, *Appl. Phys. Lett.* **85**, 378 (2004).
24. P. Clapp, in *Nanostructured Materials*, C. C. Koch, Ed. (Noyes, Norwich, NY, 2001), chap. 6.
25. I. Szlufarska, R. K. Kalia, A. Nakano, P. Vashishta, *Appl. Phys. Lett.* **86**, 021915 (2005).
26. A. Hasnaoui, H. V. Swygenhoven, P. M. Derlet, *Phys. Rev. B* **66**, 184112 (2002).
27. H. Hahn, K. A. Padmanabhan, *Nanostruct. Mater.* **6**, 191 (1995).
28. P. Heitjans, S. Indris, *J. Phys. Condens. Matter* **15**, R1257 (2003).
29. A. Hasnaoui, P. M. Derlet, H. V. Swygenhoven, *Acta Mater.* **52**, 2251 (2004).
30. J. Li, K. v. Vliet, T. Zhu, S. Yip, S. Suresh, *Nature* **418**, 307 (2002).
31. A. Gouldstone, H.-J. Koh, K.-Y. Zeng, A. E. Giannakopoulos, S. Suresh, *Acta Mater.* **48**, 2277 (2000).

32. W. W. Gerberich *et al.*, *Acta Mater.* **47**, 4115 (1999).
33. A. M. Minor, J. W. Morris, E. A. Stach, *Appl. Phys. Lett.* **79**, 1625 (2001).
34. M. Chen *et al.*, *Science* **300**, 1275 (2003).
35. W. M. Vetter, M. Dudley, *J. Cryst. Growth* **260**, 201 (2004).
36. A. T. Blumenuy *et al.*, *J. Phys. Condens. Matter* **14**, 12741 (2002).
37. J. P. Rino *et al.*, *Phys. Rev. B* **70**, 045207 (2004).
38. M. Ohyanagi *et al.*, *Scripta Mater.* **50**, 111 (2004).
39. I. A. Ovid'ko, *Science* **295**, 2386 (2002).
40. D. Wolf, V. Yamakov, S. R. Phillpot, A. Mukherjee, H. Gleiter, *Acta Mater.* **53**, 1 (2005).
41. We thank R. Kalia, D. Morgan, and S. Babcock for fruitful discussions and H. van Swygenhoven for valuable comments on the study. This work was partially supported by Air Force Office of Scientific Research–Defense University Research Initiative on Nanotechnology, Army Research Office Multidisciplinary University Research Initiative, Defense Advanced Research Projects Agency–Predicting Real Optimized Materials, the U.S. Department of Energy,

and NSF. I.S. is supported by NSF grant DMR-0512228.

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5736/911/DC1
Materials and Methods
Fig. S1
References

4 May 2005; accepted 6 July 2005
10.1126/science.1114411

Characterization of Excess Electrons in Water-Cluster Anions by Quantum Simulations

László Turi,¹ Wen-Shyan Sheu,² Peter J. Rossky^{3*}

Water-cluster anions can serve as a bridge to understand the transition from gaseous species to the bulk hydrated electron. However, debate continues regarding how the excess electron is bound in $(\text{H}_2\text{O})_n^-$ as an interior, bulklike, or surface electronic state. To address the uncertainty, the properties of $(\text{H}_2\text{O})_n^-$ clusters with 20 to 200 water molecules have been evaluated by mixed quantum-classical simulations. The theory reproduces every observed energetic, spectral, and structural trend with cluster size that is seen in experimental photoelectron and optical absorption spectra. More important, surface states and interior states each manifest a characteristic signature in the simulation data. The results strongly support assignment of surface-bound electronic states to the water-cluster anions in published experimental studies thus far.

Clusters are widely studied, both for their direct role in atmospheric and interstellar chemistry and for their intermediacy between gaseous and condensed phases, which renders them useful simplifying models for complex molecular processes in solution. Negatively charged water clusters have long been used as models to understand the hydrated electron in bulk water. Since its discovery in 1962 (*1*), the hydrated electron has been the subject of numerous experimental (2–5) and theoretical (6–10) studies for its wide-ranging role in chemical and biological electron transfer. A consistent physical picture of its structural, spectral, and dynamic properties has emerged, bolstered in part by details extracted from clusters (11–26). However, a key issue remaining with regard to the cluster data is whether the electron is bulklike, trapped in the cluster interior by oriented solvent molecules, or stabilized in a surface-bound state specific to the cluster environment. This issue bears critically on the relation of cluster observations to bulk properties and the transition from one regime to the other.

Here, we address the question through mixed quantum-classical molecular simulation, which allows the direct computation of the experimental observables for these clusters. We show that the available experimental energetic and spectral data are completely consistent with the conclusion that the anionic water clusters observed to date bind the excess electron on the surface, although the long-anticipated spontaneous transition to interior states is indicated for clusters in the range of 100 to 200 molecules.

Barnett *et al.* first identified surface states through a series of quantum mechanical simulations of negatively charged water clusters (11). For their model, they found that clusters comprising approximately 8 to 32 water molecules bind the excess electron preferentially in a localized state on the cluster surface. The calculations predicted transition to compact hydrated electron-like interior states with increasing cluster size ($32 < n < 64$). These observations parallel the later theoretical discovery by Berkowitz and co-workers (27) that polarizable atomic anions preferentially adopt surface states in clusters as well (27–29).

Experiments have provided indirect insight into the electronic structure. The comprehensive studies of photoelectron spectra in cluster-size-selected molecular beams by Coe *et al.* (15) led to an excellent correlation of the most probable vertical detachment energy [(VDE), the energy needed to remove an electron at the anion's geometry] with the cluster size, n ,

through the largest cluster measured, $n \leq 69$. For clusters of $n \geq 11$, the spectroscopic data fit well to a simple linear relationship in $n^{-1/3}$ for the size dependence, based on a dielectric model assuming interior states (11). Because the correlation line extrapolated to a value for the infinite cluster that was consistent with simulation of bulk solvated electrons in ambient water, the authors concluded that these clusters were consistent with hydrated electron-like interior states.

However, in an important theoretical work, Makov and Nitzan developed a continuum dielectric model to evaluate the energy and free-energy differences between solvation of a spherical ion (or electron) in the center versus on the surface of a spherical solvent cluster and also estimated the VDEs (13). For an ion of constant radius in a solvent with high dielectric constant, they showed that the free energy of transfer between the surface and interior of the cluster essentially vanishes. The VDE of a surface state for a negative ion was actually found to be slightly larger than for an ion at the center of the solvent shell. We note that for an electron that is expectedly more delocalized at the surface than in the interior, this difference should be compensated (or possibly overcompensated, thus reversing the VDE ordering). In addition, Makov and Nitzan showed that both interior and surface states manifest the linear scaling of the VDE with $n^{-1/3}$ seen experimentally, so that this scaling did not distinguish the excess electron-binding morphologies (13). Of particular importance, for the infinite cluster, both surface and interior states will therefore extrapolate to the same bulk limit. Hence, the experimental observation of an extrapolated value close to the bulk does not a priori distinguish between surface and interior states.

Later, Ayotte and Johnson (16) measured cluster-size-selected optical absorption spectra by photodestruction. The spectral peak positions also shift linearly with $n^{-1/3}$, consistent with an excited-state energy that also scales with cluster radius. The authors noted that the excited-state VDE slope was implicitly smaller than the ground-state slope, a result that would be in harmony with different radii for the excited and ground states. The energy gap between the ground and the excited states increased with cluster size, in accord with a contracting radius. They pointed out that, of the earlier simulated energies (11), those for interior states were quantitatively closer to the

¹Eötvös Loránd University, Department of Physical Chemistry, Budapest 112, Post Office Box 32, H-1518, Hungary. ²Department of Chemistry, Fu-Jen Catholic University, Taipei, Taiwan 242, ROC. ³Department of Chemistry and Biochemistry, Institute for Theoretical Chemistry, University of Texas at Austin, Austin, TX 78712–1167, USA.

*To whom correspondence should be addressed.
E-mail: rossky@mail.utexas.edu

measurements than those for surface states. Further, the spectra exhibited not only a blue shift with increasing size but also an increasing line width. The sequence of spectra appeared qualitatively consistent with a systematic evolution toward the bulk ambient spectrum.

Bartels summarized the entire controversy (19) and reevaluated the optical absorption spectra acquired by Ayotte and Johnson (16) based on dispersion relationships. Statistical moment analysis of the spectra yielding values for both the thermally averaged spatial dispersion $\langle r^2 \rangle$ and the kinetic energy $\langle T \rangle$ of the excess electron leads to a distinctive behavior with cluster size: The radius decreases strongly with increasing n , approaching the bulk value from above, while, in parallel, the kinetic energy approaches the bulk value from below. These quantities vary smoothly with n , without the discontinuity in either quantity that might indicate a transition between surface and interior states.

The latest work in the field has come from Neumark and co-workers (22), who measured the photoelectron spectra of larger water-cluster anions ($n \leq 150$) for clusters generated with both low and high backing pressures, the latter yielding presumably colder clusters. This study raised the question of temperature dependence directly, although Barnett and co-workers had provided a limited consideration (11) and Johnson had noted that cluster temperature is a function of preparation method and cluster size (16). We note that the uniform continuum model has only a weak temperature dependence. Neumark's group observed a new feature in the photoelectron spectra in colder clusters, with significantly smaller VDEs than those found by Coe *et al.* (15). Because of the smaller VDE, they attributed the new peaks to the presence of surface states and concluded that the earlier work had, therefore, observed interior states.

Here, we report a series of mixed quantum-classical, fully molecular simulations on $(\text{H}_2\text{O})_n^-$ clusters with $n = 20, 30, 45, 66, 104$, and 200 and with internal kinetic energies consistent with three different temperatures, 100 K, 200 K, and 300 K. The simulation methods are described in more detail elsewhere (6, 8, 10, 30). The water molecules are described classically, interacting through a flexible three-site potential, whereas the electron is represented by its wave function in a plane-wave basis evenly distributed on a 32- by 32- by 32-point grid. The water-electron interaction is modeled by a recent approximate pseudopotential model (10). Sampling is done by molecular dynamics. The water molecules evolve under the combined influence of other classical molecules and the electron. The nuclear evolution is adiabatic; the electron remains in its ground state. All the cluster simulations have been initiated from interior states of previous hydrated electron simulations (10). The equilibration of the systems has included tests to specifically establish that the surface states result

from spontaneous migration from an interior state to the cluster surface state.

In this work, we have not attempted to evaluate the relative free energies of surface and interior states. Thus, in the ensuing discussion, "stability" refers simply to the persistence of that state over the duration of the simulation. The data reported here have been obtained from equilibrated trajectories of 30- to 80-ps length. These durations are approximately an order of magnitude longer than those of Barnett and co-workers (11). The present method employs several approximations, including the use of classical nuclei, the neglect of explicit solvent polarizability, and the use of an approximate electron-water pseudopotential, which neglects electron-solvent dispersion interactions (23). Simulations using the same set of approximations for the bulk hydrated electron give results that are consistent with experiments (10). Because these approximations are expected to introduce significant quantitative error for smaller clusters (23), we consider only $n \geq 20$ here. In particular, the use of the present water model's fixed charge polarization, appropriate to the bulk liquid environment, is likely to overestimate the molecular dipole in a small cluster and, correspondingly, artificially enhance the electrostatic electron binding. Spectral results on the hydrated electron at low bulk density and high temperature (31) indicate that models like those used here are correspondingly limited in that regime.

The electronic distributions for the ground-state anions fall in two distinct categories. Identification follows from the radius of the cluster r_c , the electron radius [radius of gyration, $r_e = \langle r^2 \rangle^{1/2}$], and the distance between the centers of the electron and water distributions R . An interior state is localized within the cluster, so that $R + r_e < r_c$. For a surface state, $R \sim r_c$, and significant electronic amplitude appears outside the nominal cluster radius (Fig. 1). It is notable that the surface electronic states are highly analogous to bulk hydrated electron distributions (6), with typical s and p character. At 200 K and 300 K, we find that

only the $n = 200$ cluster exhibits a stable interior electronic state; smaller clusters ($n = 20, 30, 45, 66$, and 104) support stable surface states. At 100 K, only the smallest clusters, up to $n = 45$, spontaneously manifest the excess electron on the surface. Larger clusters, and an alternative $n = 45$ configuration, are stable with the electron in an interior state. However, this may be a kinetic effect, and we cannot conclude from this that the lower temperature favors interior states.

The validity of these calculations can be tested with the computed physical properties of the clusters, such as absorption spectra at different temperatures (Fig. 2). The spectral evolution of the surface-state clusters, in terms of both shift and width, completely parallels the experimental spectra of Ayotte and Johnson (16). There is only weak temperature dependence for surface states. The experimental spectral trends should therefore be substantially the same over a very wide range of cluster temperatures, as long as the electron is surface bound.

In contrast, the spectrum at 300 K for the $n = 200$ interior state is nearly identical to the bulk hydrated electron spectrum (10). At 200 K, the corresponding $n = 200$ spectrum is slightly blue-shifted from the bulk peak center at 298 K (10). At 100 K, the spectral evolution exhibits a sharp shift at the point of surface to interior transition (at $n \approx 45$), and at $n = 200$ is blue-shifted by nearly 0.5 eV from the bulk simulated spectrum at 298 K (10). This characteristic blue-shift from the bulk spectrum would be an experimental signature of cold interior states. We attribute the temperature dependence of the interior state clusters to solvent contraction (electrostriction) around the electron with decreasing temperature. For surface states, a contraction of solvent does not increase confinement of the electron. For the bulk hydrated electron, simulations (32) with the same model used here (10) have qualitatively reproduced the experimental temperature dependence of spectra in the liquid and shown that at liquid densities, this dependence lies predominantly, although not entirely (31), in the solvent density response.

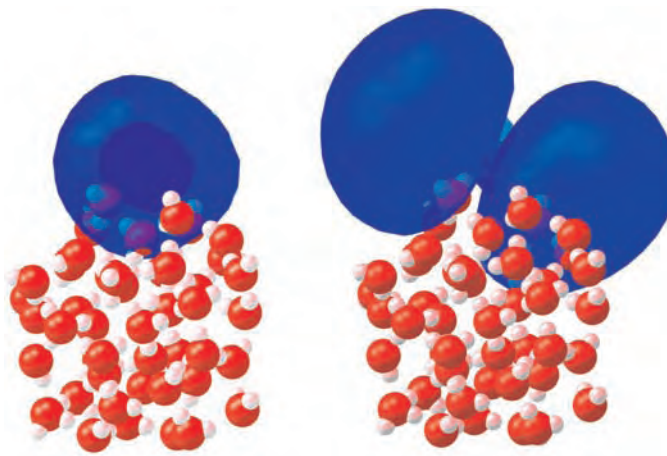


Fig. 1. Typical electronic distributions for the surface state of $(\text{H}_2\text{O})_{45}^-$. (Left) Ground state, isosurface enclosing 80% of electron density, with inner shadow isosurface at 30%. (Right) Excited state, isosurface enclosing 80% of electron density.

The temperature dependence observed here contradicts the conclusions in the seminal simulations of Barnett *et al.* (11, 12); the discrepancy is likely due to the difference in model and/or the limited sampling accessible in the older work. Our low-temperature spectra do appear qualitatively similar to those of Barnett *et al.* (12) except for the size dependence. As noted, the surface to interior transition occurred there between $n = 32$ and 64 at 300 K (11), whereas our model predicts a transition in the $104 < n < 200$ range.

The calculated spectra shown in Fig. 2 broaden and shift to the blue with increasing cluster size, in excellent agreement with experiments. The blue shift results from a sharper increase in stabilization of the s state than of the more diffuse p state as the cluster grows. The broadening is assignable primarily to an increasing p -state splitting. For smaller clusters, the p states are more nearly degenerate in each configuration, leading to overall narrower spectra.

Extrapolation of the calculated spectral maxima energies and vertical detachment energies to infinite cluster size is compared in Fig. 3 with the extrapolations of analogous experimental data from Coe (15) and Johnson (16). Because the surface states exhibit only weak

temperature dependence, the surface-state points in the plot are averaged over the temperatures simulated. The simulated surface-state data follow a slope similar to the experiments. Also, as in experiments (16), the ground-state energy (VDE) slope is different from that for the excited state, and, correspondingly, both lines extrapolate to very near the ambient bulk values for the model. This result is expected based on the continuum dielectric model (13). In contrast, the simulated interior-state data clearly differ from experimental data, and the low-temperature data do not extrapolate to room-temperature bulk properties. If the surface-state to interior-state transition does occur above $n = 104$, we find that the observed results of Coe (15) and Johnson (16) would then be consistent with our simulated results over a wide range of actual cluster temperatures.

Comparison of the calculated radius of gyration and kinetic energy data with experimental data makes an even more compelling case than the spectral data for predominance of surface states (Fig. 4). The distinct trends with cluster size follow the experimentally derived size dependence (19) faithfully [see figure 1 in (19)]. The radius smoothly approaches the bulk value from above as the clusters grow, whereas the

interior states have nearly identical radii. A similar pattern holds for the kinetic energy; the trend follows experimentally derived size dependence (19). The interior states have much higher kinetic energy than the surface states. There is some increase with decreasing temperature, but all interior-state kinetic energies are similar to the large cluster limit. These trends are qualitatively similar to those seen originally by Barnett (11), although different in magnitude.

Some additional insight into the regularity of the behavior of the surface-state clusters can be obtained by computing the mean dipole polarization, $\langle\mu\rangle$, of the molecular clusters in the direction of the electronic center of charge. The surface states manifest a variation of $\langle\mu\rangle$ in the 15 to 33 Debye range, for $20 < n < 104$ cluster sizes at 200 K, with a nearly linear dependence on $n^{-1/3}$, in accord with the expected size dependence of the Makov-Nitzan dielectric model (13). This molecular-cluster dipole moment largely compensates the dipole associated with the position of the surface-bound electron with respect to the cluster center of mass.

There are quantitative shortcomings in the calculated values compared with experimental reports. The calculated VDE values are closer to those only recently measured by Neumark (22)

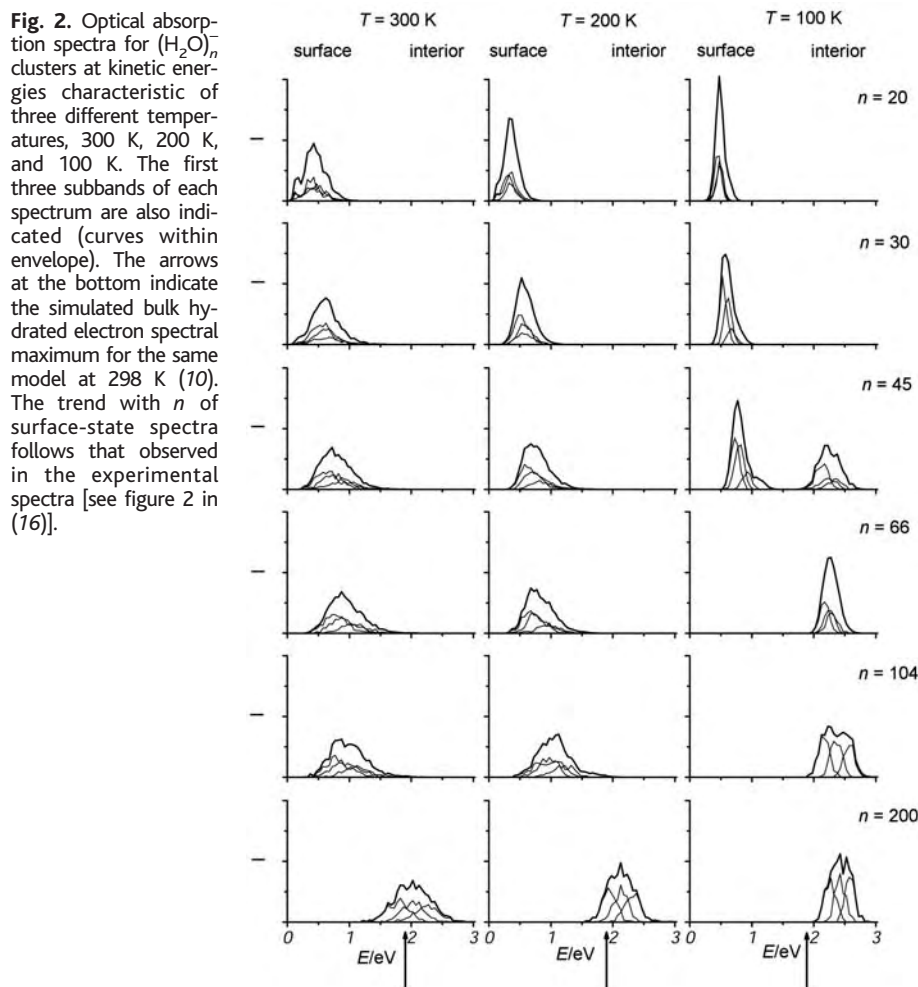


Fig. 2. Optical absorption spectra for $(\text{H}_2\text{O})_n^-$ clusters at kinetic energies characteristic of three different temperatures, 300 K, 200 K, and 100 K. The first three subbands of each spectrum are also indicated (curves within envelope). The arrows at the bottom indicate the simulated bulk hydrated electron spectral maximum for the same model at 298 K (10). The trend with n of surface-state spectra follows that observed in the experimental spectra [see figure 2 in (16)].

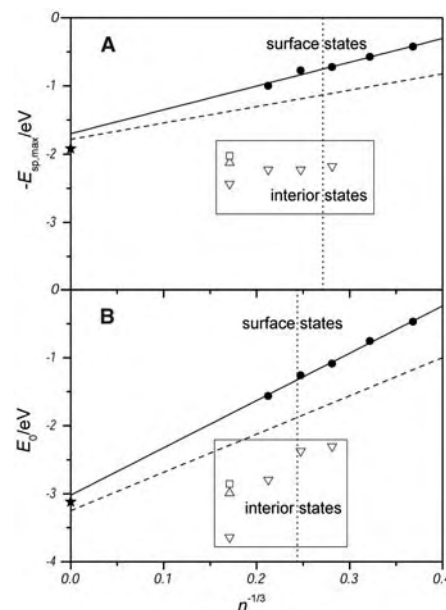


Fig. 3. Simulated spectral maxima for optical absorption (A) and mean VDEs (B) for $(\text{H}_2\text{O})_n^-$. The surface-state data (filled circles) are reported as the average over all temperatures studied. All data for the interior states (open symbols) fall in the highlighted rectangular area. ∇ , 100 K; Δ , 200 K; \square , 300 K. The bold stars at $n^{-1/3} = 0$ indicate the simulated bulk hydrated electron values at 298 K (10). The linear fits derived from the VDE data from Coe and Bowen (15) and the spectral maxima of Ayotte and Johnson (16) are also shown (dashed lines). The vertical dotted lines indicate the maximum experimental cluster sizes reported in (15) and (16). The linear extrapolation of the simulations is very similar to that seen in experiments [see figure 3 in (16)].

for the identified surface states (denoted there as Isomer II) than to the Coe data (15) considered here. However, the surface and interior electron-binding morphologies lead to distinctly different trends in measured physical properties: vertical detachment energy, optical absorption spectra, kinetic energy, and electronic radius. The comparison of the trends to the corresponding published experimental data strongly supports the conclusion that the available experiments reporting these results reflect only clusters characterized by electronic surface states. The newly reported species associated with more weakly bound electrons (22) are presumably also surface states, as concluded in that report, but they do not appear to be a simple extrapolation of those found here. We have carried out a set of preliminary simulations that show that electrons attached to already formed very cold water clusters produce surface-state species with a range of vertical electronic detachment energies of magnitudes well below those of the clusters simulated here. We therefore speculate that the surface states recently observed are the result of such attachment processes. The differences may reflect alternative proton-ordering motifs, as suggested by Johnson and co-workers (25, 26).

Our findings substantially support the physical picture originally put forth by Barnett and co-workers (11) and strongly suggest that for larger clusters than observed to date, the transition to an interior state should occur. Most important, the results indicate that both the physical state and cluster temperature of interior

states can be characterized from optical spectra or from the character of the high-energy end (VDE > 3.0 eV) of photoelectron spectra.

These results reinforce the conclusion that simple continuum dielectric models of these clusters have considerable power, particularly for surface states, but are limiting in describing the temperature dependence of the spectra and kinetic energy, key factors in the interpretation of data for interior states. The weak temperature sensitivity of surface states clearly explains why extrapolation of the surface-state properties leads to a value close to the bulk ambient VDE, because this bulk state is nearly isoenergetic with the actual extrapolation limit, the semi-infinite solvent surface state (13). The relatively large temperature dependence of interior-state properties also emerges as a convenient way to distinguish the two binding morphologies. Finally, we note that nuclear quantum effects on water structure will play a role in the quantitative comparison of experiments and simulations. Classical water clusters are expected to exhibit the characteristics of significantly colder systems when considered at the same nominal temperature as experimental (quantized) water clusters (8, 33).

References and Notes

- E. J. Hart, J. W. Boag, *J. Am. Chem. Soc.* **84**, 4090 (1962).
- A. Migus, Y. Gauduel, J. L. Martin, A. Antonetti, *Phys. Rev. Lett.* **58**, 1559 (1987).
- F. H. Long, H. Lu, K. B. Eisenthal, *Phys. Rev. Lett.* **64**, 1469 (1990).
- C. Silva, P. K. Walhout, K. Yokoyama, P. F. Barbara, *Phys. Rev. Lett.* **80**, 1086 (1998).
- M. J. Tauber, R. A. Mathies, *J. Am. Chem. Soc.* **125**, 1394 (2003).
- P. J. Rossky, J. Schnitker, *J. Phys. Chem.* **92**, 4277 (1988).
- A. Wallqvist, G. Martyna, B. J. Berne, *J. Phys. Chem.* **92**, 1721 (1988).
- B. J. Schwartz, P. J. Rossky, *J. Chem. Phys.* **101**, 6902 (1994).
- A. Staib, D. Borgis, *J. Chem. Phys.* **103**, 2642 (1995).
- L. Turi, D. Borgis, *J. Chem. Phys.* **117**, 6186 (2002).

- R. N. Barnett, U. Landman, C. L. Cleveland, J. Jortner, *J. Chem. Phys.* **88**, 4429 (1988).
- R. N. Barnett, U. Landman, G. Makov, A. Nitzan, *J. Chem. Phys.* **93**, 6226 (1990).
- G. Makov, A. Nitzan, *J. Phys. Chem.* **98**, 3459 (1994).
- H. Haberland, H. G. Schindler, D. R. Worsnop, *Ber. Bunsenges. Phys. Chem.* **88**, 270 (1984).
- J. V. Coe et al., *J. Chem. Phys.* **92**, 3980 (1990).
- P. Ayotte, M. A. Johnson, *J. Chem. Phys.* **106**, 811 (1997).
- P. Ayotte et al., *J. Chem. Phys.* **110**, 6268 (1999).
- J. V. Coe, *Int. Rev. Phys. Chem.* **20**, 33 (2001).
- D. M. Bartels, *J. Chem. Phys.* **115**, 4404 (2001).
- A. E. Bragg, J. R. Verlet, A. Kammrath, O. Cheshnovsky, D. M. Neumark, *Science* **306**, 669 (2004).
- D. H. Paik, I.-R. Lee, D.-S. Yang, J. S. Baskin, A. H. Zewail, *Science* **306**, 672 (2004).
- J. R. Verlet, A. E. Bragg, A. Kammrath, O. Cheshnovsky, D. M. Neumark, *Science* **307**, 93 (2005).
- F. Wang, K. D. Jordan, *Annu. Rev. Phys. Chem.* **54**, 367 (2003).
- K. D. Jordan, *Science* **306**, 618 (2004).
- N. I. Hammer et al., *Science* **306**, 675 (2004).
- N. I. Hammer, J. R. Roscioli, M. A. Johnson, *J. Phys. Chem. A*, in press.
- L. S. Sremaniak, L. Perera, M. L. Berkowitz, *J. Phys. Chem.* **100**, 1350 (1996).
- G. H. Peslherbe, B. M. Ladanyi, J. T. Hynes, *Chem. Phys.* **258**, 201 (2000).
- P. Jungwirth, D. J. Tobias, *J. Phys. Chem. B* **106**, 6361 (2002).
- F. A. Webster, P. J. Rossky, R. A. Friesner, *Comput. Phys. Commun.* **63**, 494 (1991).
- D. M. Bartels, K. Takahashi, J. A. Cline, T. W. Marin, C. D. Jonah, *J. Phys. Chem. A* **109**, 1299 (2005).
- C. Nicolas, A. Boutin, B. Lévy, D. Borgis, *J. Chem. Phys.* **118**, 9689 (2003).
- H. Gai, G. K. Schenter, L. X. Dang, B. C. Garrett, *J. Chem. Phys.* **105**, 8835 (1996).
- This work was supported by research grants to L.T. from the Bolyai Research Fellowship, the Foundation for Hungarian Higher Education and Research, Ministry of Education, Hungary (0140/2001), and from the National Research Fund of Hungary (OTKA, T049715); by grants to P.J.R. from the U.S. National Science Foundation (CHE 0134775) and the R. A. Welch Foundation (F-0019); and by a grant to W.S.S. from the National Science Council, ROC (NSC 93-2113-M-030-006). We thank C. Lopez for his assistance in preparing Fig. 1.

7 June 2005; accepted 8 July 2005

10.1126/science.1115808

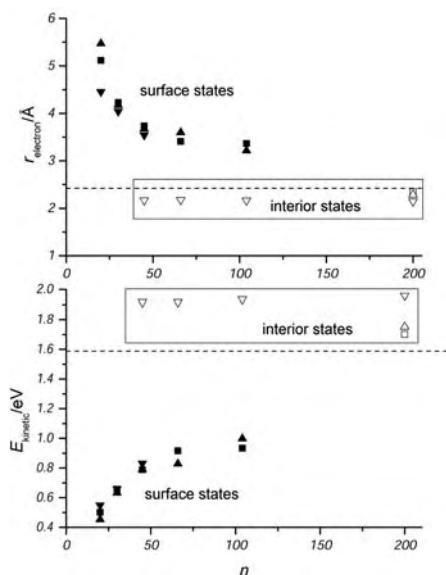


Fig. 4. Radius of gyration and kinetic energy of the excess electron in water-cluster anions at three simulation temperatures, 100 K (∇), 200 K (\square), and 300 K (\circ). Filled symbols denote surface states, open symbols denote interior states. The dashed lines show the simulated radius of gyration and kinetic energy of the hydrated electron in bulk water at 298 K (10). The surface-state data behave comparably to the experimental data [see figure 1 in (19)].

Probing the Ultrafast Charge Translocation of Photoexcited Retinal in Bacteriorhodopsin

S. Schenkl,¹ F. van Mourik,¹ G. van der Zwan,²
S. Haacke,^{1*} M. Chergui^{1,†}

The ultrafast evolution of the electric field within bacteriorhodopsin was measured by monitoring the absorption changes of a tryptophan residue after excitation of retinal. The Trp absorption decreases within the first 200 femtoseconds and then recovers on time scales typical for retinal isomerization and vibrational relaxation. A model of excitonic coupling between retinal and tryptophans shows that the signal reflects a gradual rise of the retinal difference dipole moment, which precedes and probably drives isomerization. The results suggest an intimate connection between the progressive dipole moment change and the retinal skeletal changes reported over the same time scale.

In photobiology, light-induced redistributions of charge determine the dynamic force fields at all time scales and drive the structural changes

needed for biological function (1). Photoisomerization in retinal proteins is an example in this respect. Franck-Condon excitation of all-

trans retinal results in an immediate change of permanent dipole moment by more than 12 D (2, 3), causing a sudden polarization of retinal. It has been suggested that the dielectric response of the protein might be the primary light-induced event (4, 5) that could drive structural changes on longer time scales (6). Relative to solvents, the specific environment of the protein enhances the yield of isomerized retinal and ensures stereoselectivity of the photoproduct (7), which, in the case of bacteriorhodopsin (bR), is the 13-cis isomer (8). Recently, Groma *et al.* (9) reported a macroscopic polarization of oriented bR films on a time scale of ~ 50 fs. González-Luque *et al.* (10) computed the distribution of charges in the S_1 excited state as a function of the isomerization coordinate of model chromophores of rhodopsin and bR. The computed dipole moment of the S_1 state increases along the isomerization coordinate of the chromophore, confirming an earlier suggestion by Salem and Bruckman (11).

Transient absorption studies show identical dynamics during the first 200 fs for bR and for its mutants, with locked chromophores, that cannot undergo isomerization (12). Thus, it appears that no isomerization occurs on this time scale. Transient absorption studies on wild-type bR using 5-fs pulses (13) support this conclusion and suggest that a skeletal change occurs in the first 200 fs, which corresponds to a large twisting about the C13-C14 bond. Isomerization occurs later (≤ 500 fs) and is followed by a slower vibrational relaxation of the isomerized photoproduct (the so-called J to K transition) (14–17). The time scale of charge translocation and its interplay with the initial twisting (13) and subsequent isomerization are still unknown. To address this issue, we used the natural tryptophan (Trp) residues of the protein to probe the evolution of the electric field therein after excitation of retinal.

Four tryptophans are in the vicinity of retinal in the binding pocket of bR (Fig. 1A) (18, 19): Trp⁸⁶ and Trp¹⁸² sandwich retinal, whereas Trp¹³⁸ and Trp¹⁸⁹ are located in the vicinity of the β -ionone ring. The absorption spectrum of bR (Fig. 1B) exhibits a band in the visible, due to retinal, and a band near 280 nm, predominantly due to the eight Trps present in the protein (20). The Trp absorption is due to transition from the ground state to the lowest two close-lying excited states (labeled L_a and

L_b) of the indole moiety. The transition into the L_a state implies a large difference dipole moment with respect to the ground state (21). Trp residues are therefore particularly well-suited molecular-level sensors of electric field changes within the protein, as their spectral features will be affected by these changes. In particular, Trp⁸⁶ is closest to retinal, and its difference dipole moment is almost parallel with respect to the retinal backbone (Fig. 1A, inset), so that it is particularly sensitive to the Coulomb force fields close to retinal. Note that a ~ 10 D dipole on retinal creates a field of ~ 10 MV/cm on the nearby amino acids. To probe the electric field changes associated with the charge translocation in retinal (22), we excited the latter at 560 nm and interrogated the Trp absorption changes with time resolution of 80 to 90 fs, using tunable near-UV (ultraviolet) pulses (23). Our approach has two advantages: (i) In contrast to experiments in the visible or infrared (IR) that probe retinal directly (13–16), the charge displacements become detectable with high sensitivity through an electric dipole interaction in the UV; and (ii) the local electric field changes are followed in real time from femtoseconds to picoseconds or more.

Fig. 2A shows the transient absorption changes detected between 265 and 280 nm. The traces are normalized, highlighting their similar shape over the whole range of probe wavelengths. The presence of a rise time is directly deduced from the temporal derivative of the bleach transient (Fig. 2B). Its full width at half maximum (FWHM) of 150 fs is larger

than the pump-probe cross-correlation (80 to 90 fs). For longer times, we observe a biexponential recovery of the absorption with time scales $\tau_1 = 420$ fs and $\tau_2 = 3.5$ ps; such values have been reported in experiments that probe retinal directly (14–16). A constant, weak bleach signal is also observed at the longest delay times (τ_c). The relative amplitudes of these three components are 0.72, 0.14, and 0.14, respectively.

We attribute the bleach transients of Fig. 2A to the response of Trp residues for the following reasons: (i) The transients are independent of probe wavelength over more than 2000 cm^{-1} (265 to 280 nm), indicating that a single species is responsible for the bleach. (ii) The wavelength dependence of the bleach amplitude maps reasonably well onto the Trp L_a absorption spectrum (Fig. 2C). (iii) A retinal bleach signal would follow the instrument response (14) without the additional rise time seen in Fig. 2B.

We believe that the spectral response is mainly due to the Trp⁸⁶ residue. First, the linear polarization anisotropy of the signal is high (0.30 ± 0.05), meaning that the responding Trp transition dipole is almost parallel to the transition dipole moment of retinal along the conjugate chain, as is the case for the L_a state of Trp⁸⁶ (Fig. 1A, inset). Second, we performed measurements on the W182F mutant, where Trp¹⁸² is replaced by phenylalanine, which is spectroscopically silent in the observation window. The temporal and spectral (Fig. 2C) characteristics of the signal are identical to those found in wild-type bR, pointing to the temporal evolution of Trp⁸⁶ being governed by

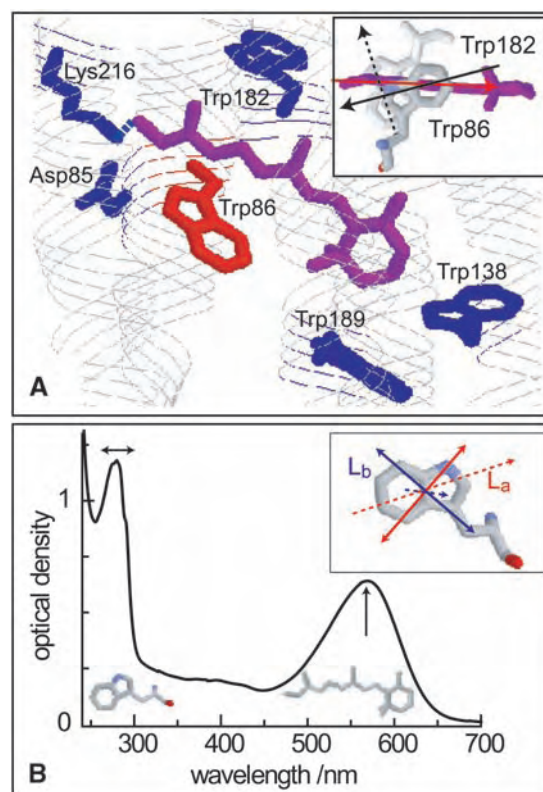


Fig. 1. Structure and optical properties of bR. (A) The retinal binding pocket of bR (from PDB entry 1C3W) (19). Retinal (purple) is covalently bound to Lys²¹⁶ through a Schiff base linkage. The four nearest Trp residues are shown. Inset: Relative orientations of the difference dipole moments for retinal (red arrow), Trp⁸⁶ (solid black arrow), and Trp¹⁸² (dashed black arrow). (B) Absorption spectrum of bR in the purple membrane. The band with maximum at 568 nm is due to the S_0 - S_1 transition of the protonated Schiff base form of retinal, whereas the near-UV band at 280 nm is dominated by the S_0 - L_a and S_0 - L_b transitions of eight Trps. The vertical and horizontal arrows indicate the excitation wavelength and the range of probe wavelengths, respectively. Inset: Orientations of the transition (solid arrows) and difference (dashed arrows) dipole moments of tryptophan for S_0 - L_b (red) and S_0 - L_a (blue) (21). Arrows point to the positive end of the difference dipole moment.

¹Ecole Polytechnique Fédérale de Lausanne, Laboratory of Ultrafast Spectroscopy, Institut de Sciences et d'Ingénierie Chimiques, FSB-BSP, CH-1015 Lausanne-Dorigny, Switzerland. ²Vrije Universiteit Amsterdam, Analytical Chemistry and Applied Spectroscopy, Faculty of Sciences, de Boelelaan 1083, 1081 HV Amsterdam, Netherlands.

*Present address: Institut de Physique et Chimie des Matériaux de Strasbourg (IPCMS), Groupe d'Optique Non-linéaire et d'Optoelectronique, 23 Rue du Loess, F-67034 Strasbourg Cédex, France.

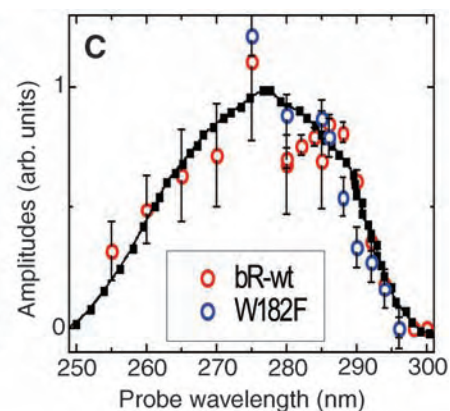
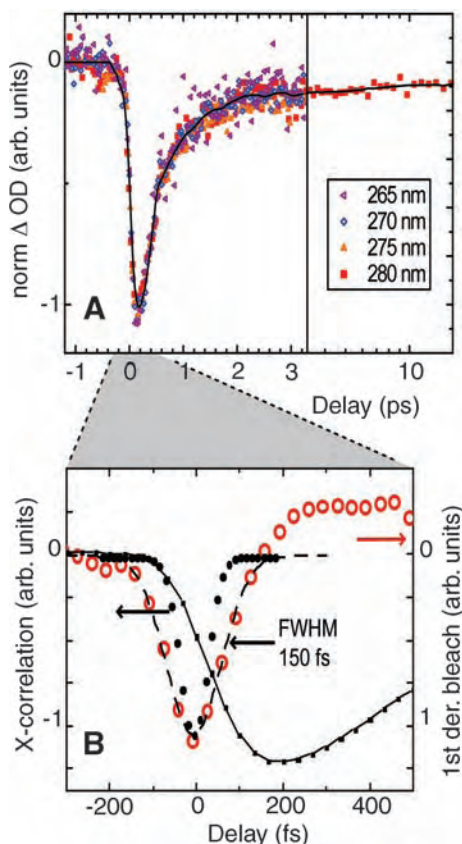
†To whom correspondence should be addressed. E-mail: majed.chergui@epfl.ch

retinal. Surprisingly, though, the amplitude of the signal increased by a factor of ~ 1.8 in the mutant. This is rationalized by the model presented in (23) and discussed below.

A dielectric response of Trps through reorientation of their dipoles is unlikely to account for the rise time, as the polarization anisotropy is constant throughout the time scale of the transients. In addition, the fast inertial relaxation of indole-like molecules is slower, even in liquids (24). Thus, the rise time (Fig. 2B) reflects both the instantaneous dipole change and the ensuing charge translocation along retinal, implying that the latter is gradual. The Trp response does not resemble a difference spectrum arising from a pure Stark shift of the L_a transition, as the bleach component does not have a positive counterpart with a comparable amplitude at nearby wavelengths. Rather, the difference spectrum is dominated by the effects of excitonic coupling resulting from the resonance interaction of retinal and Trp at ~ 280 nm, meaning that oscillator strength borrowing can occur between retinal and the Trp chromophores. To model the signal, we considered a system consisting of three coupled chromophores having three electronic levels each: S_0 , S_1 , and S_n for retinal and S_0 , L_a , and

L_b for Trp⁸⁶ and Trp¹⁸² (23). They form an excitonic complex whose proper electronic states are linear combinations of the singly excited near-UV states of $|L_{a,b}, 0, 0\rangle$, $|0, L_{a,b}, 0\rangle$, and $|0, 0, S_n\rangle$ (product basis notation $|\text{Trp}^{86}, \text{Trp}^{182}, \text{retinal}\rangle$), of which three are important (X_1 through X_3 ; Fig. 3) as a result of favorable relative orientations of the transition dipole moments. The doubly excited states (XX_1 and XX_2), with retinal in the S_1 state, arise from linear combinations of $|L_{a,b}, 0, S_1\rangle$ and $|0, L_{a,b}, S_1\rangle$ (23).

The simulated absorption changes are plotted in Fig. 3 for various values of the retinal difference dipole moment $\Delta\mu$. The photoinduced signal in the region of Trp absorption indeed appears as a bleach of the excitonic transitions into the X_i levels (the bleach character arises from the depletion of ground-state retinal as a result of excitation with the pump pulse), whereas transitions from $|0, 0, S_1\rangle$ to XX_i levels appear as a positive signal on the red side. Because the retinal difference dipole moment $\Delta\mu$ acts only when the system is in the $|0, 0, S_1\rangle$ state, the bleach ΔA is independent of $\Delta\mu$, whereas the position of the induced absorption ΔA^* red-shifts with increasing difference dipole moment. The red shifting is due



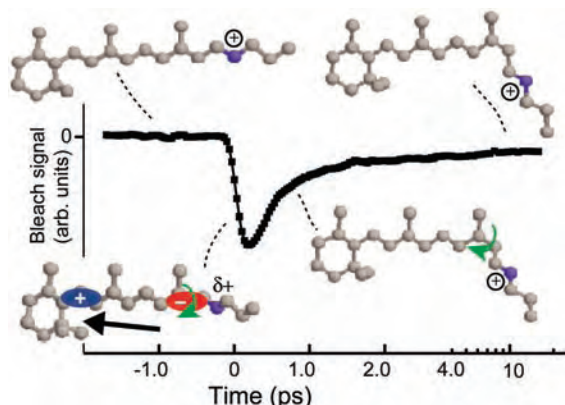
experimental time resolution of 85 fs (black dots) determined by the pump-probe cross-correlation (23). (C) Maximum amplitude of the bleach transients of wild-type bR and of the W182F mutant as a function of probe wavelength. Experimental data (circles) are compared with the L_a absorption component of Trp in propylene glycol at -50°C (25) (black squares). The L_a absorption was shifted by 5 nm to the blue to account for the apolar character of the binding pocket. The data have been normalized to the Trp L_a spectrum at 285 nm.

to an attractive interaction between the excited-state dipole moments of Trp⁸⁶ and retinal (Fig. 1, inset), with the latter varying in time. The signal is the sum of ΔA and ΔA^* , and it results in an increasing bleach signal with increasing $\Delta\mu$ (Fig. 3), as observed.

Repeating the calculation without Trp¹⁸² leads to the same $\Delta\mu$ -dependent bleach signal but with an amplitude increased by a factor of 1.75, in good agreement with our experimental results on the W182F mutant. These calculations confirm that Trp⁸⁶ is the sensor for the ultrafast charge translocation in retinal in bR. The amplitude of its signal is, however, determined by the presence of Trp¹⁸² (and most probably of the other residues) in a nonadditive fashion, which is generally true for interactions involving polarizability, as is the case here because of the inclusion of transition dipole moments.

Fig. 3. Top panel: Calculated difference absorption spectra as a function of the retinal difference dipole moment $\Delta\mu$ and its decomposition (scheme on the right) in terms of a bleach contribution ($\Delta A < 0$) and a $\Delta\mu$ -dependent absorption ($\Delta A^* > 0$). Bottom panel: Dipole-dipole coupling of Trp⁸⁶, Trp¹⁸², and retinal leads to the formation of an excitonic complex with proper electronic states (23). The X_i s represent singly excited states of the excitonic complex, with one excitation per chromophore; the XX_i s represent doubly excited states with retinal in the S_1 state. Photoexcitation of retinal attenuates the transition to the $X_1 \dots X_3$ states ($\Delta A < 0$) because of the bleach of the $|0, 0, S_0\rangle$ (product basis notation $|\text{Trp}^{86}, \text{Trp}^{182}, \text{retinal}\rangle$) ground state of the excitonic complex. The photoinduced absorption from $|0, 0, S_1\rangle$ reaches the XX_1 and XX_2 levels (in which both Trp and retinal are excited), giving rise to an absorption ($\Delta A^* > 0$). As the difference dipole moment $\Delta\mu$ increases, ΔA^* red-shifts because of a lowering of the XX_1 and XX_2 levels (attractive interaction between retinal and Trp⁸⁶). This leads to an increasing bleach around 280 nm, which is a measure of the retinal difference dipole moment $\Delta\mu$.

Fig. 4. Schematic representation of the evolution of the difference dipole moment $\Delta\mu$ and structure of retinal in bR. This scheme is based on the present results (black squares) and on literature results. Retinal and a portion of the Lys²¹⁶ side chain are represented. At $t < 0$, retinal is in the all-trans configuration and the Schiff base (blue) is fully protonated. At $0 < t < 200$ fs, preceding isomerization, a large $\Delta\mu$ (arrow) builds up as a result of charge translocation in the excited state (indicated by colored ellipses and the lower charge on the Schiff base). At $200 \text{ fs} < t < 500$ fs, isomerization brings retinal into the 13-cis ground state and the difference dipole moment decreases. At $t > 500$ fs, vibrational relaxation planarizes the 13-cis retinal, resulting in a further decrease of the difference dipole moment.



On the basis of the observed rise time of the bleach, and the calculations that link the amplitude of the bleach signal to the retinal difference dipole moment (Fig. 3), we suggest that after the sudden dipole moment change upon excitation in the Franck-Condon region at $t = 0$ (2, 3) there is an additional progressive increase of $\Delta\mu$ on a time scale of 200 fs. The increase of the bleach signal (Fig. 2) is a signature of the dipole moment change, predicted by the theoretical calculations (10), and it identifies its time scale. Given that the events at $t < 200$ fs are also characterized by structural changes of the retinal skeleton before isomerization (13), we conclude that the dipole moment increase is related to these structural changes and possibly drives the accompanying torsional motion (Fig. 4).

At later times (400 to 500 fs), the difference dipole moment decreases. This is the time scale of formation of the 13-cis isomer (16), pointing to a weaker dipolar interaction in the ground-state 13-cis retinal than in excited all-trans retinal (Fig. 4). Because the excited-state surface has mainly an ionic character (10) and the ground-state surface a mainly covalent character, the transition from one surface to the other leads to a decrease of the dipole moment. The additional 3.5-ps decay component, which is similar to the vibrational relaxation in the 13-cis photoproduct found in experiments probing retinal (14, 15, 17), suggests a further decrease of the dipolar interaction. Vibrational relaxation leads to a weak displacement of the center of gravity of the charge, which may cause the decrease in dipole strength. Alternatively, energy dissipation may cause structural changes in the retinal binding pocket that modify the field detected at the location of Trp⁸⁶. Finally, the weak bleach component at longer times (τ_{∞}) points to long-lived changes in the structure of the pocket and/or the dielectric constant, which may relate to changes detected in the picosecond range in mid-IR experiments (16). Our observations allow us to establish the connection between

the translocation of charge (10) and the skeletal changes of the conjugate chain (13).

References and Notes

1. A. Warshel, *Acc. Chem. Res.* **14**, 284 (1981).
2. R. Mathies, L. Stryer, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 2169 (1976).
3. J. Huang, Z. Chen, A. Lewis, *J. Phys. Chem.* **93**, 3314 (1989).
4. D. Xu, C. Martin, K. Schulten, *Biophys. J.* **70**, 453 (1996).
5. J. T. M. Kennis et al., *J. Phys. Chem. B* **106**, 6067 (2002).
6. A. Lewis, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 549 (1978).
7. G. Wald, *Science* **162**, 230 (1968).
8. M. Tsuda, M. Glaccum, B. Nelson, T. G. Ebrey, *Nature* **287**, 351 (1980).
9. G. I. Groma et al., *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7971 (2004).
10. R. González-Luque et al., *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9379 (2000).
11. L. Salem, P. Bruckmann, *Nature* **258**, 526 (1975).

12. Q. Zhong et al., *J. Am. Chem. Soc.* **118**, 12828 (1996).
13. T. Kobayashi, T. Saito, H. Ohtani, *Nature* **414**, 531 (2001).
14. R. A. Mathies, C. H. B. Cruz, W. T. Pollard, C. V. Shank, *Science* **240**, 777 (1988).
15. J. Döbler, W. Zinth, W. Kaiser, D. Oesterhelt, *Chem. Phys. Lett.* **144**, 215 (1988).
16. J. Herbst, K. Heyne, R. Diller, *Science* **297**, 822 (2002).
17. S. J. Doig, P. J. Reid, R. A. Mathies, *J. Phys. Chem.* **95**, 6372 (1991).
18. E. Pebay-Peyroula, G. Rummel, J. P. Rosenbusch, E. M. Landau, *Science* **277**, 1676 (1997).
19. H. Luecke, B. Schobert, H.-T. Richter, J.-P. Cartailler, J. K. Lanyi, *J. Mol. Biol.* **291**, 899 (1999).
20. A possible contribution of tyrosines can be neglected, as their extinction coefficient is much smaller than for Trps.
21. P. R. Callis, *Methods Enzymol.* **278**, 113 (1997).
22. The positive charge from the protonated Schiff base linkage with Lys²¹⁶ (Fig. 1) is partially moved toward the ionone ring.
23. See supporting data on Science Online.
24. M. L. Horng, J. A. Gardecki, A. Papayian, M. Maroncelli, *J. Phys. Chem.* **99**, 17311 (1995).
25. M. R. Eftink, L. A. Selvidge, P. R. Callis, A. A. Rehms, *J. Phys. Chem.* **94**, 3469 (1990).
26. Supported by Swiss National Science Foundation contracts 21-55830.98 and 2153-65135.01 and the National Center for Competence in Research-Quantum Photonics. We thank N. Friedman and M. Sheves for the purple membrane samples and for insightful discussions; R. Schlesinger and G. Büldt for the W182F mutant; J. Bredenbeck for the design of the flow cell; and M. Olivucci for critical reading of the manuscript.

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5736/917/DC1

Materials and Methods

Figs. S1 and S2

Tables S1 and S2

References

24 February 2005; accepted 21 June 2005
10.1126/science.1111482

An Experimental Approach to the Percolation of Sticky Nanotubes

B. Vigolo, C. Coulon, M. Maugey, C. Zakri, P. Poulin*

Percolation is a statistical concept that describes the formation of an infinite cluster of connected particles or pathways. Lowering the percolation threshold is a critical issue to achieve light and low-cost conductive composites made of an insulating matrix loaded with conductive particles. This has interest for applications where charge dissipation and electrical conductivity are sought in films, coatings, paints, or composite materials. One route to decreasing the loading required for percolation is to use rod-like particles. Theoretical predictions indicate that this may also be achieved by altering the interaction potential between the particles. Although percolation may not always respond monotonically to interactions, the use of adhesive rods can be expected to be an ideal combination. By using a system made of carbon nanotubes in an aqueous surfactant solution, we find that very small attraction can markedly lower the percolation threshold. The strength of this effect can thereby have direct technological interest and explain the large variability of experimental results in the literature dealing with the electrical behavior of composites loaded with conducting rods.

Percolation is a statistical concept that describes the formation of an infinite cluster of connected particles or pathways (1). This concept is widely used to study various phenomena such as flow in porous media and conducting-insulating transitions in compos-

ites. Percolation is driven by thermodynamics and occurs above a given concentration. At this threshold, a composite made of an insulating matrix loaded with conducting particles becomes conductive and can be used for antistatic or electromagnetic shielding applica-

tions. Lowering the percolation threshold is a critical issue to reduce the cost and facilitate the processing of such composites. A route to this goal uses rod-like particles. These particles exhibit a larger excluded volume than spheres of the same weight. As a result, they form percolated networks at concentrations lower than spherical particles (2, 3). This is why conductive rod-like fillers, chopped fibers, nanowires, and nanotubes are currently widely investigated for making antistatic and electrically conductive materials (4–12). It may be possible to further decrease the percolation threshold by making the particle interact (13–16). Although this effect is not systematically monotonic as a function of the strength and range of the interactions (15), Wang and Chatterjee recently predicted that the percolation threshold of rod-like particles should slightly decrease if the particles become attractive (16). In spite of numerous theoretical studies, mostly for spheres, percolation of adhesive rods remains elusive from a quantitative and experimental point of view.

We present experimental results and a simple analytical model that provide a quantitative approach to this problem. In contrast with the few available predictions, we found a very strong effect. We show that a variation as small as $0.006kT$ (where k is Boltzmann's constant and T is temperature) in the interaction potential between perpendicular rods allows the percolation threshold to be decreased by a factor of 3. We believe that the exceptional strength of this effect can shed light on the large variability of experimental results dealing with composites loaded with conducting rods (4–12).

We used bundles of HiPco (17, 18) single-wall carbon nanotubes (CNTs) dispersed in water and stabilized by sodium dodecyl sulfate (SDS), an ionic surfactant. CNTs typically form bundles with an average diameter of a few nm and a length of 1 μm (19, 20). They are used as model electronic conductive rods. SDS molecules form micelles, which induce a depletion attraction between the nanotubes (21, 22). The strength of the attraction is directly proportional to the osmotic pressure of the micelles, which is itself proportional to the micelle concentration. Attractive interactions between the rods can thereby be accurately controlled by varying the surfactant concentration. Interaction differences in distinct samples can be as small as $10^{-4}kT$. The polar head groups of the SDS molecules are dissociated in water; thus, the solvent is a conductive electrolyte. This prevents the determination of the percolation threshold of the CNTs by direct measurements of the real part of the conductivity. For this characterization, an insulating organic matrix would be more

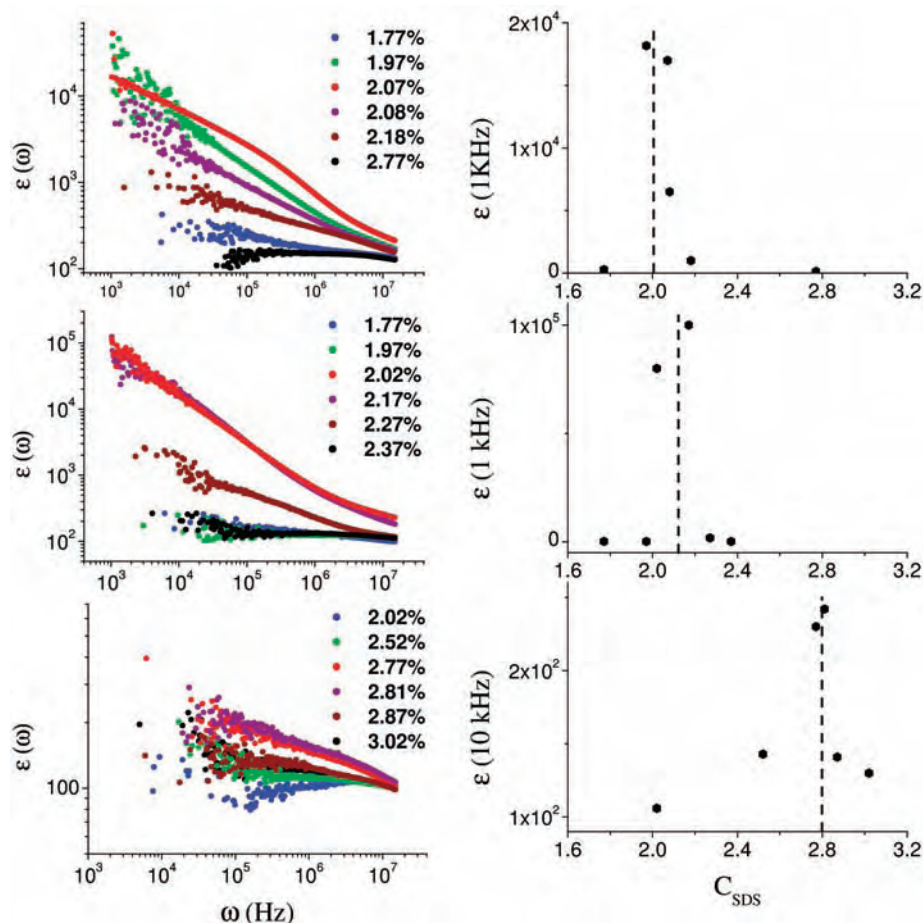


Fig. 1. (Left) Dielectric constant of nanotube dispersions versus frequency for different micelle weight fractions, indicated by different colors. The nanotube volume fractions are, from top to bottom, 0.19, 0.13, and 0.06%. (Right) A determination of the surfactant concentration corresponding to the percolation threshold by identifying the points where the dielectric constant at low frequency is maximum. The percolation is indicated by the dashed lines.

suitable. However, it would be more difficult to achieve an accurate control of the dispersion and of the interaction potential in such a system. Nevertheless, the nanotubes and the solvent exhibit a dielectric contrast with different charge carriers. This is why percolation can still be detected via dielectric measurements. The formation of large polarizable clusters, as the percolation threshold is approached, yields an increase of ϵ_0 , the dielectric constant at zero frequency (8, 23). Above the percolation, the clusters are connected between the electrodes and can no longer be polarized, and ϵ_0 strongly decreases. Consequently, the percolation is expected to be revealed by a divergence of ϵ_0 . Measurements at zero frequency are experimentally difficult because of double-layer polarization effects at the particle and electrode interfaces (24, 25). Nevertheless, the increase at zero frequency is so pronounced that the percolation threshold can still be determined by measurements at somewhat larger frequencies, in the kHz range. In this range, a sharp peak is observed in place of an actual divergence. We refer to the values we measured as ϵ_{IP} , the low-frequency dielectric constant. Considering the

sharpness of the peak, this approach remains sufficient for our main goal, which is the determination of the percolation threshold. Furthermore, CNTs strongly absorb light, and this allows a qualitative assessment of their aggregation behavior by optical microscopy.

CNT bundles in a surfactant solution form clusters when the surfactant concentration is increased (21, 22). Their presence can be revealed by optical micrographs of dispersions with the same CNT fraction and increasing SDS concentration (26). When the SDS concentration is increased, the aggregates become larger and more evident. Qualitatively, the amount of surfactant required to form clusters is larger when the CNTs are diluted. These features are reminiscent of a phase separation controlled by the competition between the entropy of the dispersion and the depletion attraction. However, we stress that the system does not undergo a phase separation. Instead, for a given CNT concentration, a gradual evolution takes place with the continuous growth of clusters as the SDS concentration is increased. The structures are stable and do not coarsen with time, and the aggregates reversibly redisperse upon dilution.

Centre de Recherche Paul Pascal, CNRS, Avenue Schweitzer, 33600 Pessac, France.

*To whom correspondence should be addressed. E-mail: poulin@crrp-bordeaux.cnrs.fr

They can also be destroyed upon shear and spontaneously reform at rest. The present features are fundamentally different from a kinetic gelation, as observed in other colloids (27) out of equilibrium and where the adhesion is much greater than kT . Here, the system is a distribution of thermodynamically stable clusters with an average size that increases when the rods experience stronger attractions. As a consequence, a percolated network can form above a certain surfactant concentration.

Dielectric measurements confirm that the growth of the clusters leads to percolation (24, 25). Not initially knowing the threshold, we tested a series of samples by keeping the CNT fraction constant and by slightly varying the amount of SDS. The dielectric responses of different samples are shown in Fig. 1. A huge increase of the dielectric constant at low frequency is observed in a narrow SDS concentration range (28). The increase seen in the data confirms the percolation and allows the threshold to be determined. We examined three nanotube volume fractions—0.06%, 0.13%, and 0.19%—for a series of SDS concentrations (18). Experiments outside this range have proven to be difficult. When the CNT concentration is too low, the dielectric response is weak, and its accurate measurement is more difficult. When the CNT volume fraction is too high, typically above 0.2%, it is experimentally difficult to obtain homogeneous dispersions. A

possible origin of this limitation may indeed be the percolation presently studied. Figure 1 shows only the concentration range where a sharp increase of the dielectric constant is observed. Above and below the shown SDS concentration range, the dielectric constant of all the tested samples is close to that of the solvent. The sharp variation of ϵ_{if} allows an accurate determination of the SDS micelle weight fraction at the percolation threshold. The results are summarized in Fig. 2, where the black symbols represent the percolation threshold as function of the micelle weight fraction (28).

To quantitatively account for these results, we derive the percolation threshold following approaches that consider the theory of pair connectedness to describe the distribution of physical clusters of particles (13, 29, 30). In the limit of dilute systems, the percolation threshold is expected to be inversely proportional to the integral of $\exp(-\beta u^+)$, where u^+ is the interaction potential of so-called bound particles and $\beta = 1/kT$. The definition of bound particles is somewhat arbitrary but has to have its basis in physical grounds (30). The actual potential u between the bundles is approximated to a sticky potential plus a hard core repulsion. As another approximation, we assume that these bundles are monodispersed with a square section of effective width d and effective length L (Fig. 3). In this context, a reasonable definition for bound particles consists of choosing $u^+ = u$ when the bundles are in contact; that is, when the distance between their long axes is d . We choose $u^+ = +\infty$ for larger or smaller separations. The latter describes the absence of overlapping particles because of the hard core repulsion. These approximations smear out the details of the interaction potential, which is difficult to model considering the actual polydispersity in shape

and size of the CNT bundles. A different definition for the potential and for the connectivity criteria would lead to a determination of different effective dimensions of the bundles. Nevertheless, because the interaction range is on the order of d and the anisotropy ratio L/d is very large, other reasonable choices would not change the main conclusions and orders of magnitude of the effective dimensions. Therefore, we propose to take for the percolation threshold the following integral over γ , the angular, and r , the spatial, coordinates:

$$\varphi_p = \frac{4Ld^2}{\int_V \int_0^\pi \exp(-\beta u) d^3r \sin(\gamma) d\gamma} \quad (1)$$

As shown in Fig. 3, γ is the angle between the rods. The prefactor is chosen so that $\phi_p \rightarrow \phi_0 = d/L$, the percolation threshold in the absence of interactions (2) when $u \rightarrow 0$. An exact calculation of Eq. 1 is not easily tractable. Nevertheless, following (31) and (32) where the second Virial coefficient of interacting rods is calculated, we solve Eq. 1 by separating the integral into two terms. The first term corresponds to the situation where the rods make an angle γ greater than γ_c , where γ_c is given by $\sin \gamma_c = \frac{d}{L}$. The second term corresponds to the situation of almost parallel rods for γ smaller than γ_c . The potential u is of the form $u = \frac{w_\perp}{\sin \gamma}$ when the rods are in contact and at an angle γ , as illustrated in Fig. 3. w_\perp is the adhesion energy between perpendicular rods that are in contact. When the rods are in contact and parallel, the interaction potential is given by $u = \frac{w_\perp}{d} z$, where z is the contact length. With these assumptions, the integral in Eq. 1 is

$$2L^2 d \int_{\gamma_c}^{\frac{\pi}{2}} \exp\left(\frac{\beta w_\perp}{\sin \gamma}\right) \sin(\gamma) d\gamma + 4d^2 \int_0^L \exp\left(\frac{\beta w_\perp}{d} z\right) dz \quad (2)$$

which can be directly computed and the percolation threshold determined. We point out that this simple approach contains two important and universal limits. The first term in Eq. 2 is expected to be dominant for weak interactions, whereas the second prevails for strong interactions. This means that parallel contacting rods should be more frequent for strong adhesion and that the structure, as well as the geometry of contacts, should be notably different upon varying the strength of the interaction. Lastly, we further assume that w_\perp is directly proportional to C_{SDS} , the surfactant micelle weight fraction. The anisotropy ratio L/d and the proportionality factor between C_{SDS} and w_\perp are the only two independent fitting parameters. Even though the range of data are limited because of experimental constraints, the available points allow us to estimate the validity of our claims. The plot in Fig. 2 attempts to fit the data with different

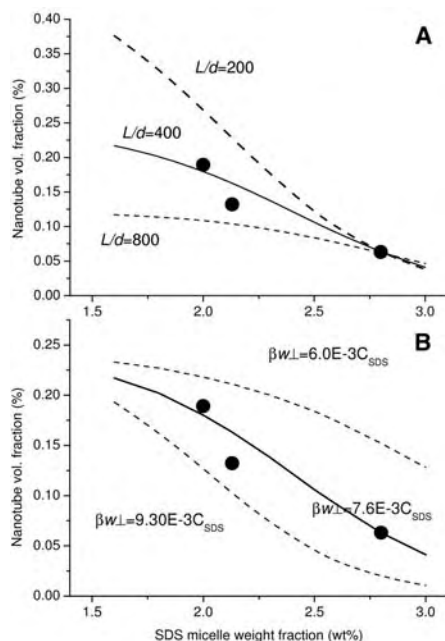


Fig. 2. Percolation threshold of interacting nanotube bundles. Black circles indicate experiments and dashed and continuous lines, percolation threshold versus surfactant concentration calculated by the model. The best fit to the data is shown by the continuous line. It leads to $L/d = 400$ and $\beta w_\perp = 7.6 \times 10^{-3} C_{SDS}$. Dashed lines in (A), $\beta w_\perp = 7.6 \times 10^{-3} C_{SDS}$ is kept constant and L/d varied. Dashed lines in (B), $L/d = 400$ is kept constant and $\beta w_\perp / C_{SDS}$ varied.

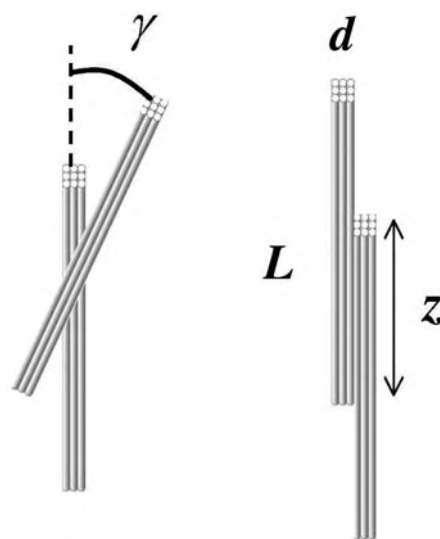


Fig. 3. Schematic of interacting nanotube bundles. (Left) The bundles are in contact at an angle γ . (Right) The bundles are in contact and parallel. L and d are the length and the width of the bundles.

sets of parameters. Keeping w_{\perp} fixed, the L/d are chosen (Fig. 2A) so that all the lines intersect the data point for the highest CNT concentration. We see that the anisotropy ratio can neither be larger than 800 nor smaller than 200. This range of anisotropy ratio is in agreement with recent measurements of the dimensions of surfactant-stabilized nanotube bundles that used dynamic light scattering (19) or atomic force microscopy (20). The proportionality factor between C_{SDS} and w_{\perp} is also limited to a restricted range. We find for the best fit $\beta w_{\perp} = 7.6 \times 10^{-3} C_{\text{SDS}}$, where C_{SDS} is given in weight % (wt%). This proportionality factor can be independently estimated and linked to the diameter of the CNT bundles. As a first approximation, w_{\perp} is given by $w_{\perp} = N_m k T d^2 D_m$ where N_m and D_m are the concentration and diameter of the SDS micelles. $N_m k T$ is the osmotic pressure of the micelles and d^2 the effective contact area of the bundles. A SDS micelle contains about 75 molecules and has a diameter of about 4 nm. We can deduce the effective width d from the proportionality factor between C_{SDS} and w_{\perp} . Taking $\beta w_{\perp} = 7.6 \times 10^{-3} C_{\text{SDS}}$, we find $d \sim 3$ nm. Considering the crudeness of this estimation, the result is reasonable and supports the concept of depletion-induced percolation in rod-like suspensions. These results help to understand why studies in other systems, in the micrometer-size range, would be significantly more difficult. Similar variations of the surfactant concentration would yield variations of thousands of kT in the interaction potential, limiting thereby experimental studies.

The reported experiments show that the percolation threshold is decreased by 300% whereas w_{\perp} varies by only $0.006kT$. However,

the net adhesion between the rods can be much stronger if the rods are stuck parallel over their entire length. The maximal adhesion would be $w_{\perp} L/d$ in that limit. Thus, in contrast to spheres, a weak variation of the conditions can have dramatic effects in rod-like dispersions. Furthermore, the predicted differences of contact and cluster structures as the interaction potential is varied provide another distinctive feature of percolating rods. They should lead to differences in the physical properties of composites loaded with interacting rods.

References and Notes

- J. P. Clerc et al., *Ann. Phys.* **8**, 3 (1983).
- I. Balberg, N. Binenbaum, N. Wagner, *Phys. Rev. Lett.* **52**, 1465 (1984).
- F. Carmona, P. Prudhon, F. Barreau, *Solid State Commun.* **51**, 255 (1984).
- S. A. Curran et al., *Adv. Mater.* **10**, 1091 (1998).
- J. N. Coleman et al., *Phys. Rev. E* **58**, R7492 (1998).
- J. Sandler et al., *Polymer* **40**, 5967 (1999).
- M. S. P. Shaffer, A. H. Windle, *Adv. Mater.* **11**, 937 (1999).
- P. Potschke, S. M. Dudkin, I. Alig, *Polymer* **44**, 5023 (2003).
- J. M. Benoit, B. Corraze, O. Chauvet, *Phys. Rev. B* **65**, 241405 (2003).
- S. Barrau, P. Demont, A. Peigney, C. Laurent, C. Lacabanne, *Macromolecules* **36**, 5187 (2003).
- O. Breuer, U. Sundararaj, *Polym. Compos.* **25**, 630 (2004).
- L. A. Hough et al., *Phys. Rev. Lett.* **93**, 168102 (2004).
- Y. C. Chiew, E. D. Glandt, *J. Phys. A Math. Nucl. Gen.* **16**, 2599 (1983).
- S. Safran, I. Webman, G. S. Grest, *Phys. Rev. A* **32**, 506 (1985).
- A. L. R. Bug et al., *Phys. Rev. Lett.* **55**, 1896 (1985).
- X. Wang, A. Chatterjee, *J. Chem. Phys.* **118**, 10787 (2003).
- P. Nikolaev et al., *Chem. Phys. Lett.* **313**, 91 (1999).
- HiPco CNTs have been supplied by Carbon Nanotechnologies Incorporated (Houston, TX). The used batch contains 15 wt% (2 atomic %) iron catalyst. The CNT volume fraction is deduced by subtracting the amount of catalyst to the sample weight and by multiplying the CNT weight fraction with 1.35 g/cm^3 , the CNT bundle density. All the samples are homogenized in the same conditions by using a sonicator probe operating at 40 W.

- S. Badaire, P. Poulin, M. Maugey, C. Zakri, *Langmuir* **20**, 10367 (2004).
- M. F. Islam, E. Rojas, D. M. Bergey, A. T. Johnson, A. G. Yodh, *Nano Lett.* **3**, 269 (2003).
- J. M. Bonard et al., *Adv. Mater.* **9**, 827 (1997).
- B. Vigolo et al., *Science* **290**, 1331 (2000).
- D. Strout, D. J. Bergman, *Phys. Rev. B* **25**, 2061 (1982).
- CNT dispersion (4 ml) is poured in a Tacussel (Villeurbanne, France) conductivity cell adapted to a HP4194A (Hewlett-Packard, Palo Alto, CA) impedance meter as described in (25). The cell is kept at constant temperature (22°C) in a water temperature-controlled bath. The low-frequency dielectric constant ϵ_{if} is taken at 10^4 or 10^3 Hz. The measurements are repeated until the response is stable and reproducible. Equilibrium is typically reached 10 min after the sample cell is filled. At longer times, the measurements are reproducible, confirming thereby the absence of coarsening.
- C. Vinches, C. Coulon, D. Roux, *J. Phys. II* **4**, 1165 (1994).
- Optical micrographs of dispersions with the same CNT fraction and increasing SDS concentration are shown in fig. S1. When the surfactant concentration is increased, the aggregates become larger and more evident. In fig. S1b, small clusters actually seem to be disconnected from each other. In contrast, in fig. S1, c and d, the systems exhibit interconnected structures.
- J. Bibette, T. G. Mason, H. Gang, D. A. Weitz, *Phys. Rev. Lett.* **69**, 981 (1992).
- C_{SDS} is the micelle weight fraction. It corresponds to the total surfactant weight fraction minus the fraction of monomers, which is equal to the critical micellar weight fraction. For the surfactant used, the critical micellar fraction is 0.23 wt%.
- A. Coniglio, U. De Angelis, A. Forlani, *J. Phys. A Math. Nucl. Gen.* **10**, 1123 (1977).
- T. L. Hill, *J. Chem. Phys.* **23**, 617 (1955).
- A. Y. Grosberg, A. R. Khokhlov, *Adv. Polym. Sci.* **41**, 53 (1981).
- P. van der Schoot, T. Odijk, *J. Chem. Phys.* **97**, 515 (1992).
- This work has been done in the framework of the Groupement de Recherche-Européen no. 2756.

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5736/920/DC1
Fig. S1

28 March 2005; accepted 5 July 2005
10.1126/science.1112835

The Pyrite-Type High-Pressure Form of Silica

Yasuhiro Kuwayama,^{1*} Kei Hirose,^{1,2} Nagayoshi Sata,²
Yasuo Ohishi³

Silica (SiO_2) exhibits extensive polymorphism at elevated pressures. X-ray diffraction measurements showed that a high-pressure form with a pyrite-type structure, denser than other known silica phases, is stable above 268 gigapascals and 1800 kelvin. The silicon coordination number increases from 6 in the $\alpha\text{-PbO}_2$ -type phase to 6+2 in the pyrite-type phase, leading to a large increase in density by about 5% at the phase transition.

The high-pressure behavior of silica has long been of great interest, owing to its wide range of implications in geophysics, materials science, and solid-state physics (1). Si is tetrahedrally coordinated by O in silica polymorphs at relatively low pressures. Stishovite (a rutile-type SiO_2 phase) forms above ~ 10 GPa with an increase in the coordination number of Si from 4 to 6. Previous experimental studies have shown that stishovite undergoes a second-order structural phase transition to a CaCl_2 -type phase

around 70 GPa and 1600 K (2–5). It further transforms to an $\alpha\text{-PbO}_2$ -type phase above 121 GPa and 2400 K (6, 7) without change in the coordination number of Si. In addition, a number of metastable silica phases with 6-coordinated Si have been reported below 120 GPa (8, 9). Recent theoretical calculations have predicted a phase transition from the $\alpha\text{-PbO}_2$ -type to a pyrite-type structure around 200 GPa (10–13). However, no experimental studies of silica have been made at such multimegabar

pressures because of substantial difficulties in both compression and heating. Here we report experimental data on the crystal structure and stability of pyrite-type silica.

Angle-dispersive x-ray diffraction (XRD) spectra were collected at beamline BL10XU of SPring-8 (Japan) (14). α -quartz and silica glass (fused silica) were used as starting material in the first and third runs and in the second run, respectively. They were mixed with fine platinum powder that served as both an internal pressure standard and a laser absorber. The sample mixture was loaded into a 20- μm hole in the rhenium gasket, together with insulation layers of pure quartz or silica glass that was not

¹Department of Earth and Planetary Sciences, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro, Tokyo 152-8551, Japan. ²Institute for Research on Earth Evolution, Japan Agency for Marine-Earth Science and Technology, 2-15 Natsushima-cho, Yokosuka, Kanagawa 237-0061, Japan. ³Japan Synchrotron Radiation Research Institute, 1-1 Koto, Mikazuki-cho, Sayo-gun, Hyogo 679-5198, Japan.

*To whom correspondence should be addressed.
E-mail: ykuwayam@geo.titech.ac.jp

mixed with platinum. They were compressed by 60- μ m culet-beveled diamond anvils, with a shape best optimized for laser-heating at pressures over 200 GPa. Heating was achieved by a focused donut-mode TEM₀₁*-mode Nd:yttrium-lithium-fluoride laser with the double-sided heating technique (15), which minimizes both radial and axial temperature gradients in the sample. Temperature was measured from one side by the spectroradiometric method (16). The uncertainty in temperature within the 20- μ m areas from which XRD spectra were collected was about ± 300 K. Pressure measurements employed the equation of state for platinum proposed by Holmes *et al.* (17), which was used in previous phase equilibria studies on SiO₂ (5, 7), with (111), (200), (220), and (311) lines. The pressure error here was about ± 2 GPa. We conducted three separate runs.

In the first run, the sample was compressed to 194 GPa at room temperature. The quartz starting material became amorphous under compression; thus, we observed diffraction peaks only from the platinum and gasket material before heating (Fig. 1A). It was subsequently heated to 1800 K for 1 hour at 194 GPa. The diffraction peaks of α -PbO₂-type silica appeared within several minutes after the temperature reached 1800 K (Fig. 1A). The XRD pattern did not change with further heating. In the second set of experiments, amorphous silica was first heated to 1400 K for 1 hour at 220 to 222 GPa. We observed similar diffraction patterns, including peaks from the α -PbO₂-type phase and platinum. The sample was further compressed to 237 GPa at room temperature and reheated to 1500 to 2000 K for another hour at 254 to 257 GPa. The diffraction pattern did not change at these pressure and temperature conditions. In the third run, the sample was first squeezed to 283 GPa at room temperature and was subsequently heated. Eight new diffraction peaks appeared within 5 min during heating at 271 GPa and 1400 K (Fig. 1B). These peaks grew with increasing temperature to 1800 K at 268 GPa. Two-dimensional XRD images obtained after heating for 1 hour showed circular Debye rings for these new peaks (Fig. 2). The peaks can be indexed by a cubic unit cell with lattice parameter $a = 3.9299(2)$ at 271 GPa and 300 K and assigned to a pyrite-type (modified fluorite-type) structure (space group $Pa\bar{3}$). Figure 1C and Table 1 show the results of profile fitting by the Rietveld method (18).

Phase relations of SiO₂ are summarized in Fig. 3. The pressure-induced phase transformation from the α -PbO₂-type to the pyrite-type structure is consistent with theoretical predictions (11–13) and previous experimental results on analog materials (19, 20). Our results show that the phase boundary is located at ~ 260 GPa and 1800 K. The pressure of phase transition is slightly higher than that determined by theory. Previous theoretical

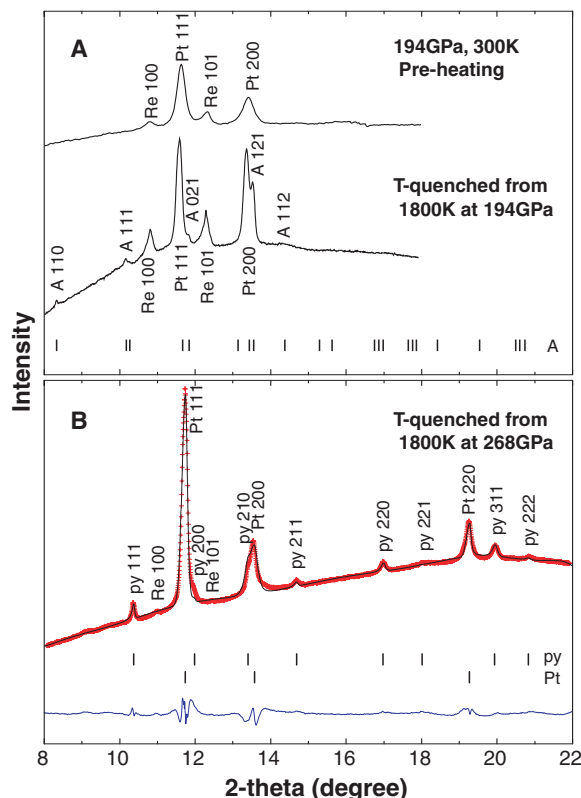


Fig. 1. XRD patterns of SiO₂. (A) Top trace: Before heating at 194 GPa and 300 K. Bottom trace: The α -PbO₂-type phase at 178 GPa and 300 K, after heating for 1 hour at 194 GPa and 1800 K. (B) Observed (red crosses) and Rietveld-fitted pattern (black line) of pyrite-type silica at 271 GPa and 300 K, after heating for 1 hour at 268 to 271 GPa and 1400 to 1800 K. A, α -PbO₂-type silica; py, pyrite-type silica; Pt, platinum; Re, rhenium (gasket). Vertical bars indicate the calculated peak positions. The difference profile (blue line) is on the same scale.

Table 1. Structural details of pyrite-type SiO₂ at 271 GPa and 300 K.

Crystal system	Cubic
Space group	$Pa\bar{3}$
Cell parameters	$a = 3.9299(2) \text{ \AA}$ $Z = 4$ $V = 60.694(6) \text{ \AA}^3$
Atomic coordinates	
Si (4a)	0, 0, 0
O (8c)	$x, x, x; x = 0.3484(11)$
Isotropic atomic displacement parameter, B_{iso} (\AA^2)	
Si	0.37(7)
O	1.48(20)
Reliability factors*	$R_{wp} = 1.38, R_p = 0.88$
Interatomic distances (\AA)	
Si–O	$1.608(4) \times 6$
Si–O†	$2.372(4) \times 2$
O–O	$2.112(6) \times 6$
O–O†	$2.063(6) \times 6$

* R_{wp} and R_p are the reliability factors of Rietveld refinement: $R_{wp} = [\sum_i w_i \{y_i(obs) - y_i(calc)\}^2 / \sum_i w_i y_i(obs)^2]^{1/2}$ and $R_p = \sum_i |y_i(obs) - y_i(calc)| / \sum_i y_i(obs)$.
†Shortest interpolyhedral atomic distance.

studies have shown that pyrite-type silica is stable above 226 GPa at 0 K (11), 205 GPa at 0 K (12), and 200 GPa at 1800 K (13).

The crystal structure of pyrite-type silica is illustrated in Fig. 4. The pyrite structure has unusual 6+2 cation coordination. Rietveld analysis showed that pyrite-type silica has six Si–O bonds with 1.608 \AA distance and an additional two interpolyhedral Si–O bonds with 2.372 \AA distance (Table 1). Such interpolyhedral Si–O distance in the pyrite-type structure is much

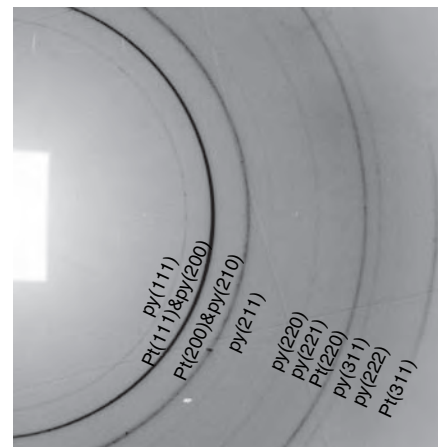


Fig. 2. Two-dimensional XRD image of pyrite-type silica at 271 GPa and 300 K.

shorter than that in the α -PbO₂-type structure (2.726 \AA when the average Si–O distance in SiO₆-octahedra is 1.649 \AA) (12), in which Si is 6-coordinated. The increase in the coordination number of Si from 6 to 6+2 results in a large increase in density (Fig. 4). The density of pyrite-type silica is 6.576(1) g/cm³ at 271 GPa and 300 K. It is larger by 4.7% than that of the α -PbO₂-type phase when compared at equivalent pressure. FeS₂ pyrite has relatively short S–S covalent bonds and is the cause of its metallic nature. If O–O bonds exist in pyrite-type silica, it possibly indicates a pressure-induced metallization of SiO₂. However, the Rietveld analysis showed that the shortest O–O

Fig. 3. Phase diagram of SiO_2 . Open circles and solid squares indicate the stabilities of the α - PbO_2 -type and pyrite-type phases, respectively. Solid reversed triangles and open triangles show previous experimental data (7) on the stabilities of the CaCl_2 -type and α - PbO_2 -type phases, respectively. The dotted lines show the tentative phase boundary.

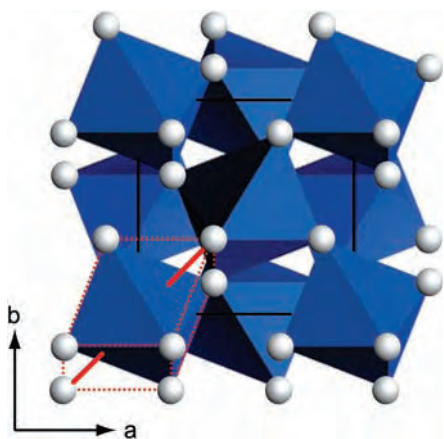
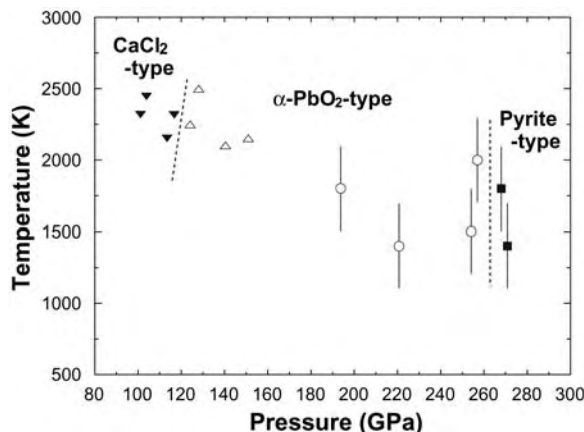


Fig. 4. Crystal structure of pyrite-type silica. Coordination polyhedra of O atoms around Si atoms are shown as octahedra. O atoms are illustrated as balls. The black line indicates the unit cell, and bold red lines show the second nearest bonds. 6+2 coordination polyhedra are illustrated by dotted red lines.

distance in pyrite-type silica is $2.063(6)$ Å (Table 1), consistent with theoretical calculations set at 0 K (11, 13), which is much longer than the typical distance of single covalent O–O bonds (~ 1.5 Å). These experimental data support the arguments that there are no O–O covalent bonds in pyrite-type silica (13).

Silica is the most abundant oxide component in the Earth's crust and mantle, but the pressure required for the pyrite type is greater than that found in the Earth's mantle. Nevertheless, it is also one of the most important oxide components in other planets of our solar system. Theoretical modelling of the interiors of ice giant planets suggests that both Uranus and Neptune may have a rocky core at pressures about 800 GPa and below (21), and therefore the pyrite-type silica might be an important constituent of these planets. Moreover, silicates are significant oxide components in extrasolar systems (22). During the formation of terrestrial planets, pressures may exceed 260 GPa and thus include the pyrite-type silica phase.

References and Notes

- R. J. Hemley, C. T. Prewitt, K. J. Kingma, in *Silica*, P. J. Heaney, C. T. Prewitt, G. V. Gibbs, Eds. (Mineralogical Society of America, Washington, DC, 1994), pp. 41–84.
- Y. Tsuchida, T. Yagi, *Nature* **340**, 217 (1989).
- K. J. Kingma, R. E. Cohen, R. J. Hemley, H. K. Mao, *Nature* **374**, 243 (1995).
- D. Andrault, G. Fiquet, F. Guyot, M. Hanfland, *Science* **282**, 720 (1998).
- S. Ono, K. Hirose, M. Murakami, M. Isshiki, *Earth Planet. Sci. Lett.* **197**, 187 (2002).
- L. S. Dubrovinsky et al., *Nature* **388**, 362 (1997).
- M. Murakami, K. Hirose, S. Ono, Y. Ohishi, *Geophys. Res. Lett.* **30**, 1207, doi:10.1029/2002GL016722 (2003).
- J. Haines, J. M. Léger, F. Gorelli, M. Hanfland, *Phys. Rev. Lett.* **87**, 155503 (2001).
- L. S. Dubrovinsky et al., *Phys. Earth Planet. Inter.* **143–144**, 231 (2004).
- K. T. Park, K. Terakura, Y. Matsui, *Nature* **336**, 670 (1988).
- B. B. Karki, M. C. Warren, L. Stixrude, G. J. Ackland, J. Crain, *Phys. Rev. B* **55**, 3465 (1997).
- D. M. Teter, R. J. Hemley, G. Kresse, J. Hafner, *Phys. Rev. Lett.* **80**, 2145 (1998).
- A. R. Oganov, M. J. Gillan, G. D. Price, *Phys. Rev. B* **71**, 064104 (2005).
- Angle-dispersive XRD spectra were obtained on an imaging plate (Rigaku R-Axis IV) and x-ray charge-coupled device (Bruker APEX) at BL10XU of SPring-8. The incident x-rays were monochromatized to a wavelength of 0.41045 or 0.41296 Å. The x-ray beam was collimated to 20 μm in diameter. Two-dimensional diffraction images were integrated as a function of 2-theta in order to give conventional one-dimensional diffraction profiles (23).
- G. Shen, H. K. Mao, R. J. Hemley, *Proceedings of the 3rd NIRIM International Symposium on Advanced Materials* (National Institute for Research in Inorganic Materials, Tsukuba, Japan, 1996), pp. 149–152.
- T. Watanuki, O. Shimomura, T. Yagi, T. Kondo, M. Isshiki, *Rev. Sci. Instrum.* **72**, 1289 (2001).
- N. C. Holmes, J. A. Moriarty, G. R. Gathers, W. J. Nellis, *J. Appl. Phys.* **66**, 2962 (1989).
- F. Izumi, T. Ikeda, *Mater. Sci. Forum* **198**, 321 (2000).
- J. Haines, J. M. Léger, O. Schulte, *Science* **271**, 629 (1996).
- S. Ono, T. Tsuchiya, K. Hirose, Y. Ohishi, *Phys. Rev. B* **68**, 014103 (2003).
- T. Guillot, *Science* **286**, 72 (1999).
- J. B. Pollack et al., *Astrophys. J.* **421**, 615 (1994).
- A. P. Hammersley, European Synchrotron Radiation Facility internal report no. ESRF97HA02T (European Synchrotron Radiation Facility, Grenoble, France, 1997).
- We thank K. Kawamura, T. Tsuchiya, H. Genda, M. Ikoma, Y. Tatsumi, T. Suzuki, and S. Ono for discussions. In situ XRD measurements were conducted at SPring-8 (proposal no. 2004B4013-LD2-np and 2005A5013-LD2-np). Y.K. was supported by the Japan Society for the Promotion of Science Research Fellowships for Young Scientists.

16 May 2005; accepted 29 June 2005
10.1126/science.1114879

Ice Sheet and Solid Earth Influences on Far-Field Sea-Level Histories

Sophie E. Bassett,¹ Glenn A. Milne,^{1*} Jerry X. Mitrovica,² Peter U. Clark³

Previous predictions of sea-level change subsequent to the last glacial maximum show significant, systematic discrepancies between observations at Tahiti, Huon Peninsula, and Sunda Shelf during Lateglacial time ($\sim 14,000$ to 9000 calibrated years before the present). We demonstrate that a model of glacial isostatic adjustment characterized by both a high-viscosity lower mantle (4×10^{22} Pa s) and a large contribution from the Antarctic ice sheet to meltwater pulse 1A (~ 15 -meters eustatic equivalent) resolves these discrepancies. This result supports arguments that an early and rapid Antarctic deglaciation contributed to a sequence of climatic events that ended the most recent glacial period of the current ice age.

The evolution of high-latitude global ice volumes, as inferred from observations of far-field sea-level change, serves as a funda-

mental constraint on ice-age climate models (1). In this regard, data from Barbados (2, 3), the Sunda Shelf (4), Tahiti (5), Huon Peninsula (6, 7) and the Bonaparte Gulf (8) (Fig. 1) record a spatially and temporally variable sea-level history that samples ice sheet fluctuations through the complex, modulating influence of glacial isostatic adjustment (GIA) (9–11). Efforts to fit these histories with global ice reconstructions and

¹Department of Earth Sciences, University of Durham, Durham DH1 3LE, UK. ²Department of Physics, University of Toronto, Toronto, Ontario M5S 1A7, Canada. ³Department of Geosciences, Oregon State University, Corvallis, OR 97331, USA.

*To whom correspondence should be addressed. E-mail: g.a.milne@durham.ac.uk

numerical models of GIA-induced sea-level change have been characterized by enigmatic and highly contentious misfits (11–14) at

several key sites, particularly during the Lateglacial period [~ 14 to 9 calibrated kiloyears before present (cal kyr BP)] (11).

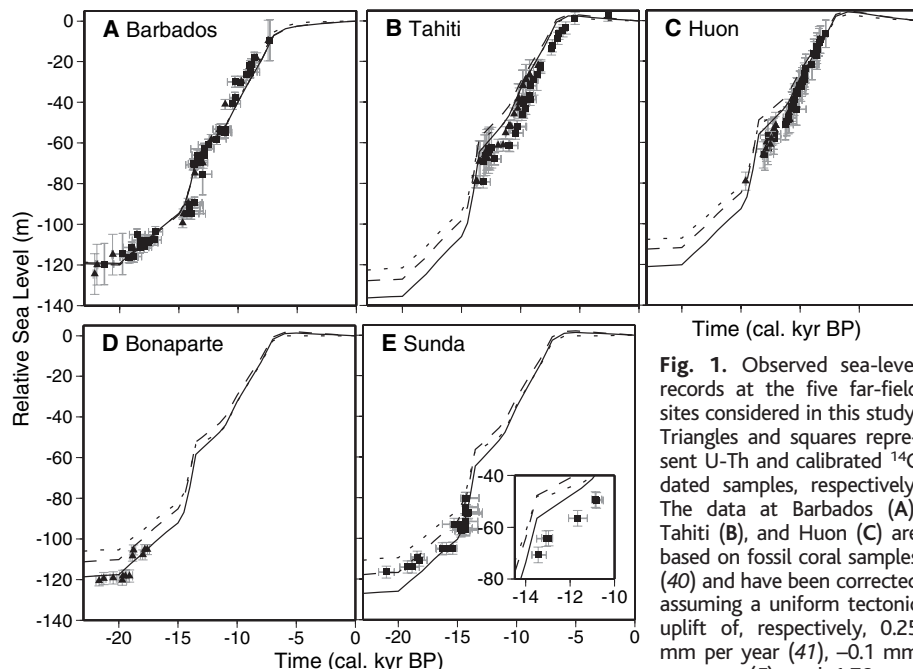


Fig. 1. Observed sea-level records at the five far-field sites considered in this study. Triangles and squares represent U-Th and calibrated ^{14}C dated samples, respectively. The data at Barbados (A), Tahiti (B), and Huon (C) are based on fossil coral samples (40) and have been corrected assuming a uniform tectonic uplift of, respectively, 0.25 mm per year (41), -0.1 mm per year (5), and 1.76 mm per year (42, 43).

The data from Bonaparte Gulf (D) and Sunda Shelf (E) are largely based on the analysis of organic material found in sediment cores. No tectonic correction has been applied to these data. Sea-level markers at Sunda Shelf are collected from two regions displaced by as much as ~ 800 km (4), and numerical predictions show significant differential trends across this zone (9). Accordingly, data (and predictions) in (E) are separated into two subsets, one from the southern zone of collection (main figure) and the other from the northern zone (inset). The lines on each panel are site-dependent numerical predictions of post-LGM sea-level change and are based on ice history that includes a dominant North American source for the mwp-IA event. The predictions are distinguished on the basis of the lower mantle viscosity of the adopted Earth model: dotted line, 2×10^{21} Pa s; dashed line, 10^{22} Pa s; and solid line, 4×10^{22} Pa s.

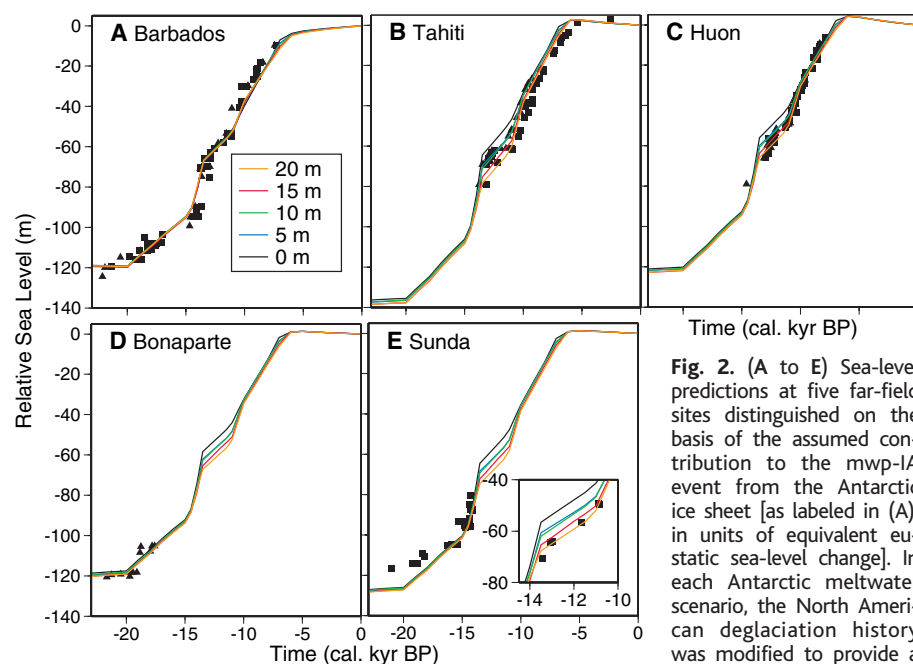


Fig. 2. (A to E) Sea-level predictions at five far-field sites distinguished on the basis of the assumed contribution to the mwp-IA event from the Antarctic ice sheet [as labeled in (A), in units of equivalent eustatic sea-level change]. In each Antarctic meltwater scenario, the North American deglaciation history was modified to provide a good fit to the Barbados

sea-level record. The predictions are based on the Earth model with $v_{\text{LM}} = 4 \times 10^{22}$ Pa s. Error bars have been omitted to improve clarity. The inset of panel (E) compares predictions with the observations derived from the more northerly subset of the cores.

Here, we investigate the origin of these Lateglacial misfits and explore whether they might be reconciled through a revision of widely adopted ice histories or Earth model parameters.

To illustrate the nature of the misfits, we superimpose GIA predictions onto the observed sea-level records in Fig. 1 that were generated by spherically symmetric, self-gravitating, and viscoelastic Earth models with density and elastic structure taken from seismic constraints (15). The Earth models are characterized by a 100-km-thick elastic lithosphere and a sublithospheric upper mantle viscosity of 5×10^{20} Pa s; the predictions are distinguished on the basis of the adopted lower mantle viscosity (below a depth of 670 km). For each choice of lower mantle viscosity (v_{LM}), the adopted ice model was modified (16) from the global ICE-3G deglaciation history (17) to fit the Barbados sea-level record. This modification involved the inclusion of a dominant North American source (18) for meltwater pulse IA (mwp-IA), as in the more recent ICE-4G and ICE-5G deglaciation models (19, 20); the mwp-IA event, first identified in the Barbados record (2, 3), resulted in a eustatic sea level rise of ~ 20 to 25 m between 14.5 and 13.5 cal kyr BP. The sea-level predictions are based on an algorithm (21) that includes an accurate treatment (22–24) of time-evolving continental shorelines and feedback from Earth-rotation perturbations.

Our focus is the ~ 5 -kyr period subsequent to the mwp-IA event (the Lateglacial period). The dotted line in Fig. 1 was generated by an ice and Earth model ($v_{\text{LM}} = 2 \times 10^{21}$ Pa s) broadly similar to several earlier GIA studies (11, 19). This ice history and Earth model combination, tuned to closely match the Barbados record, yields a poor fit to the Lateglacial sea-level histories in the Huon Peninsula, Tahiti, and Sunda Shelf (Fig. 1E, inset); in these cases, the predictions are too shallow by ~ 20 m. A similar misfit between an earlier GIA prediction (19) and data at Huon Peninsula was noted by Edwards (12), who questioned the accuracy of the GIA model; Peltier (13) suggested, in contrast, that the steady tectonic correction of 1.9 mm per year applied to the raw data (6, 12) was suspect. This argument has been reiterated in more recent work (11, 14), where it is cited to explain “extremely large” (~ 30 -m) discrepancies at both Huon and Tahiti (11). The possibility has also been raised that the misfits reflect a change in the living depth of coral assemblages from Lateglacial to early Holocene times (11, 25).

A lower mantle viscosity of 2×10^{21} Pa s is not compatible with a number of recent inferences based on data associated with GIA and/or mantle convection (26–28). Lambeck *et al.* (9) recently explored fits between GIA

predictions and far-field relative sea-level data before and after the last glacial maximum (LGM) and concluded that discrepancies over this broad time window are reduced as ν_{LM} is increased to values of $\sim 3 \times 10^{22}$ Pa s (29). The remaining lines in Fig. 1, which sample ν_{LM} values of 10^{22} Pa s (dashed line) and 4×10^{22} Pa s (solid line), support this conclusion. The discrepancies during the Lateglacial period are $\sim 50\%$ less for the latter model compared with those of the original case of $\nu_{LM} = 2 \times 10^{21}$ Pa s. Furthermore, increasing the lower mantle viscosity of the Earth model to 4×10^{22} Pa s improves the fits to the oldest Bonaparte Gulf data (9). However, this model overpredicts the LGM lowstand at Sunda Shelf by ~ 10 m. The origin of this misfit is unclear, though it has been suggested that the bulk sediment analysis used to date the oldest sections of the core at this site may have biased the ages upward by as much as several thousand years (9).

Increasing the lower mantle viscosity above 4×10^{22} Pa s produces no further improvement in the fit to the post-LGM sea-level records. This is clear from fig. S1A (dashed line), which tracks the χ^2 misfit of this sequence of post-LGM predictions as a function of ν_{LM} .

The large residual misfit evident in Fig. 1 at Tahiti, Huon, and Sunda Shelf (Fig. 1E, inset) during the Lateglacial period appears to be exacerbated by the inclusion of the mwp-IA event into the adopted ice history in an effort to reconcile the Barbados record (11). This suggests that the discrepancies might be associated with an error in the adopted ice history, namely the distribution of the total meltwater flux among the various ice complexes during the massive mwp-IA event. An obvious target of investigation is therefore the primary source region for this event, which immediately precedes the Lateglacial time window. Clark *et al.* (30) argued that the relative size of the observed sea-level jump across mwp-IA (i.e., across 14.5 to 13.5 cal kyr BP) at Barbados and Sunda Shelf precludes a sole North American source for the event. Their analysis highlighted several possible mwp-IA scenarios (31), including a potentially large Antarctic contribution to the event. This specific suggestion is supported by climate models that indicate that a freshwater flux into the Southern Ocean provides a trigger for the Bolling-Allerød warm interval (32) and by South Atlantic records of contemporaneous (to mwp-IA) ice-rafted debris originating from the Antarctic ice sheet (33).

In Fig. 2, we show a suite of predictions in which the contribution of the Antarctic ice sheet to mwp-IA is varied from 0 m (as in Fig. 1) to 20 m of equivalent eustatic sea-level rise for the Earth model in which $\nu_{LM} = 4 \times 10^{22}$ Pa s. As the magnitude of the mass flux

from the Antarctic ice sheet is increased, the contribution from the North American ice complex is suitably decreased to maintain a good fit to the Barbados record. The discrepancies between predictions and the Lateglacial sea-level trends recorded at Tahiti, Huon Peninsula, and Sunda Shelf (Fig. 1E, inset) are significantly reduced when an Antarctic component to mwp-IA is introduced.

Figure S1B shows the variation in the χ^2 misfit as this balance in meltwater source is altered. Because the variation in the predictions is most pronounced over the Lateglacial period (~ 14 to 9 cal kyr BP; Fig. 2), the misfit in fig. S1B is computed over this time window. The statistical tests directly reflect the results shown in Fig. 2: The model fit is improved to greater than 99% confidence,

compared with the dominant North American scenario, when a substantial portion (~ 5 m or more) of mwp-IA is source from Antarctica. The quality of fit for the 5- and 10-m Antarctic source models is similar. A further significant improvement in fit was obtained when the magnitude was increased to 15 or 20 m; the data from Sunda Shelf, in particular, benefit from this increase in the Antarctic contribution (Fig. 2E, inset). We adopted the 15-m Antarctic scenario, which includes ~ 8 m from northern hemisphere sources (~ 6 m from North America), as the optimal model in consideration of recent independent constraints on the volume of the LGM Antarctic ice sheet (34, 35).

The improvement in the fit is due to a number of factors. The change in the direct grav-

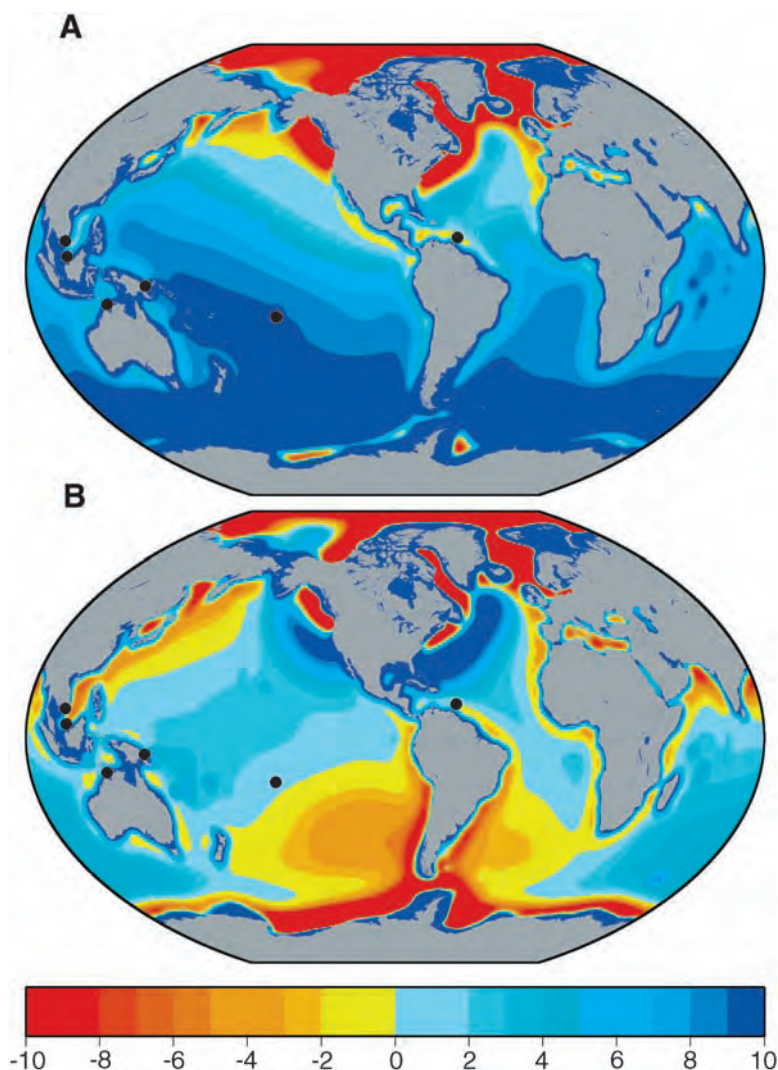


Fig. 3. Predicted sea-level change, in meters, over the past 13 cal kyr relative to the predicted change at Barbados (i.e., predictions of the raw sea-level change across this time interval are shifted by the specific prediction at Barbados, and thus the Barbados prediction falls on the zero contour in each panel). (A) The scenario in which North American ice is the sole source for the mwp-IA event (as in the predictions of Fig. 1; $\nu_{LM} = 4 \times 10^{22}$ Pa s). (B) The scenario assuming that the Antarctic ice sheet contributes an equivalent eustatic sea-level rise of 15 m to the event (as in the predictions shown by red lines in Fig. 2).

itational effect of the surface mass load as the Antarctic proportion of the mwp-IA source is increased reduces the magnitude of the sea-level rise across the event at the three southern sites (30, 36). In addition, the viscoelastic response of the planet plays an important role in the isostatic adjustment following mwp-IA. Maps of the predicted sea-level change over the past 13 cal kyr (i.e., relative sea-level change since Lateglacial time) for the scenarios in which the Antarctic ice complex contributes either 0 or 15 m to mwp-IA (Fig. 3) indicate a significant change in the geometry of the associated GIA effects. Each map is plotted relative to the prediction at Barbados. The observed relative sea level of markers of age ~13 cal kyr BP at Barbados, Tahiti, Huon Peninsula, and Sunda Shelf are all at a depth of ~60 m (Figs. 1 and 2). In the scenario where Antarctic mass flux dominates the mwp-IA event (Fig. 3B), the predicted zero contour passes close to each of these sites and thus the 13 cal kyr BP relative sea-level markers are fit by this scenario. In contrast, the dominant North American source for mwp-IA yields significantly more discrepant predictions at these four sites (Fig. 3A).

We considered two additional mwp-IA source scenarios: a global distribution of sources based on the "All-ICE-3G" scenario described in (30) and a northern hemisphere source distribution based on the "North-ICE-3G" scenario described in (30) (fig. S2). The northern hemisphere model produces a fit very similar to that generated for the case of a dominant North American source, giving a relatively high χ^2 value of 15.7 over the Lateglacial period (compare this with fig. S1B). In comparison, the global source model provides an improved fit (χ^2 of 8.5) but there remain systematic discrepancies at Huon Peninsula and Sunda Shelf that are resolved when a larger (15-m) Antarctic contribution is adopted.

We conclude that the Lateglacial far-field sea-level record is a powerful constraint for testing a range of plausible mwp-IA source scenarios and that a substantial Antarctic contribution is a robust requirement of the data.

Inspection of the Barbados sea-level record has led to suggestions of a meltwater event (mwp-IB) with onset at 11.5 cal kyr BP; however, the existence of the event has remained controversial (5, 37). To explore this issue, we altered our optimal ice model (15-m Antarctic scenario) to consider two models that include the mwp-IB event (fig. S3). The revised models produce a moderately improved fit to the early Holocene sea-level record at Barbados, but they introduce large discrepancies in the records at Tahiti and the Huon Peninsula over the same time period. Thus, our analysis does

not support the existence of the mwp-IB event.

We show that long-debated discrepancies between predictions and observations of Lateglacial sea levels at Tahiti, Huon Peninsula, and Sunda Shelf can be reconciled by adopting both a relatively high lower mantle viscosity and a large (~15-m eustatic) contribution to mwp-IA from the Antarctic ice sheet. This conclusion is contrary to previous suggestions that the discrepancy is due to uncertainties in the habitation depth of coral species or errors in the tectonic corrections applied to the raw sea-level markers (11, 13). Our results focus further attention on the Antarctic ice sheet as a key trigger for climatic events that led the Earth system out of the previous glacial period (32), and they add to our growing understanding (9, 10) of the complex space-time mapping between ice sheet ablation and global sea-level change.

References and Notes

1. K. Lambeck, T. M. Esat, E.-K. Potter, *Nature* **419**, 199 (2002).
2. R. G. Fairbanks, *Nature* **342**, 637 (1989).
3. E. Bard, B. Hamelin, R. G. Fairbanks, A. Zindler, *Nature* **345**, 405 (1990).
4. T. Hanebuth, K. Stattegger, P. M. Grootes, *Science* **288**, 1033 (2000).
5. E. Bard *et al.*, *Nature* **382**, 241 (1996).
6. J. Chappell, H. Polach, *Nature* **349**, 147 (1991).
7. K. B. Cutler *et al.*, *Earth Planet. Sci. Lett.* **206**, 253 (2003).
8. Y. Yokoyama, K. Lambeck, P. De Dekker, P. Johnston, L. K. Fifield, *Nature* **406**, 713 (2000).
9. K. Lambeck, Y. Yokoyama, T. Purcell, *Quat. Sci. Rev.* **21**, 343 (2002).
10. G. A. Milne, J. X. Mitrovica, D. P. Schrag, *Quat. Sci. Rev.* **21**, 361 (2002).
11. W. R. Peltier, *Quat. Sci. Rev.* **21**, 377 (2002).
12. R. L. Edwards, *Science* **267**, 536 (1995).
13. W. R. Peltier, *Science* **267**, 536 (1995).
14. W. R. Peltier, in *Coastal Tectonics*, I. S. Stewart, C. Vita-Finzi, Eds. (Geological Society Special Publication, London, 1998), no. 146, pp. 1–30.
15. A. M. Dziewonski, D. L. Anderson, *Phys. Earth Planet. Inter.* **25**, 297 (1981).
16. Two preliminary modifications were made to the ICE-3G deglaciation model (17): The deglaciation chronology was revised to be consistent with the ^{14}C calibration results of Bard *et al.* (3) and a glaciation phase was included by reversing the deglaciation history and extending the duration of the time increments to create a glaciation period of ~100 kyr. However, further modification to this ice history was required to fit the Barbados record. In particular, the ICE-3G model underpredicts the sea-level rise at Barbados by ~15 m (10) and does not include a rapid rise in sea level at the time of mwp-IA. We altered the North American component of the model to address these limitations. The volume of the North American ice complex was increased by ~30% (the precise value was a function of the adopted lower mantle viscosity) in accord with recent glaciological (38) and geodetic (39) constraints and a large reduction in volume was imposed at the time of mwp-IA. The duration of the ice model time increments was reduced to 0.5 cal kyr between 15 and 13 cal kyr BP to better capture this rapid event. In creating the predictions shown in Fig. 1, we made no attempt to model the mwp-IB event at ~11 cal kyr BP (2).
17. A. M. Tushingham, W. R. Peltier, *J. Geophys. Res.* **96**, 4497 (1991).
18. Specifically, the revised model included 26 m from North American ice, 2 m from other northern hemisphere sources, and no contribution from Antarctica.
19. W. R. Peltier, *Science* **265**, 195 (1994).
20. W. R. Peltier, *Annu. Rev. Earth Planet. Sci.* **32**, 111 (2004).
21. G. A. Milne, J. X. Mitrovica, J. L. Davis, *Geophys. J. Int.* **139**, 464 (1999).
22. J. X. Mitrovica, G. A. Milne, *Geophys. J. Int.* **154**, 253 (2003).
23. J. X. Mitrovica, *Quat. Sci. Rev.* **22**, 127 (2003).
24. K. Lambeck, T. Purcell, P. Johnston, M. Nakada, Y. Yokoyama, *Quat. Sci. Rev.* **22**, 309 (2003).
25. W. R. Peltier, *Rev. Geophys.* **36**, 603 (1998).
26. J. X. Mitrovica, A. M. Forte, *J. Geophys. Res.* **102**, 2751 (1997).
27. K. Lambeck, C. Smither, M. Ekman, *Geophys. J. Int.* **135**, 375 (1998).
28. G. Kaufmann, K. Lambeck, *J. Geophys. Res.* **107**, 2280 (2002).
29. Lambeck *et al.* (9) did not include an mwp-IA event in their models; although they reported systematic discrepancies during the Lateglacial period, these do not appear to be as large as those reported in studies that include the major meltwater pulse.
30. P. U. Clark, J. X. Mitrovica, G. A. Milne, M. E. Tamisiea, *Science* **295**, 2438 (2002).
31. In addition to scenarios in which the entire Antarctic ice complex, or either the east and west portions, served a dominant source for mwp-IA, Clark *et al.* (30) found that the relative size of the sea-level jump across mwp-IA at Barbados and Sunda Shelf was consistent with the following freshwater sources: the Barents Sea plus Fennoscandian ice complexes, and melting distributed over all ice sheets in the ICE-3G deglaciation model (17) (All-ICE-3G), with a weighting given by the excess volume of the various complexes at the onset of mwp-IA. A scenario in which meltwater was distributed across the northern hemisphere only (North-ICE-3G) was also considered, but it performed less well than the above scenarios.
32. A. J. Weaver, O. A. Saenko, P. U. Clark, J. X. Mitrovica, *Science* **299**, 1709 (2003).
33. S. L. Kanfoush *et al.*, *Science* **288**, 1815 (2000).
34. G. H. Denton, T. J. Hughes, *Quat. Sci. Rev.* **21**, 193 (2002).
35. P. Huybrechts, *Quat. Sci. Rev.* **21**, 203 (2002).
36. W. E. Farrell, J. A. Clark, *Geophys. J. R. Astr. Soc.* **46**, 647 (1976).
37. I. Shennan, *J. Quat. Sci.* **14**, 715 (1999).
38. S. J. Marshall, T. S. James, G. K. C. Clarke, *Quat. Sci. Rev.* **21**, 175 (2002).
39. W. R. Peltier, *J. Quat. Sci.* **17**, 491 (2002).
40. Each index point based on fossil coral samples (Barbados, Huon, and Tahiti) is plotted such that it lies at the mean depth of the habitation range of the associated coral species. The error bar reflects the range of this habitation depth. The error bars at Huon, with the exception of the most recently acquired data (7), have been increased by ± 1.5 m to account for the uncertainty in the tectonic correction for this region (9).
41. U. Radtke, R. Grün, H. P. Schwartz, *Quat. Res.* **29**, 197 (1988).
42. J. Chappell *et al.*, *Earth Planet. Sci. Lett.* **141**, 227 (1996).
43. The most recently published data from the Huon Peninsula were corrected using the rates provided by (7).
44. This research was supported by the Natural Environment Research Council of the UK (S.E.B.); the Natural Sciences and Engineering Research Council of Canada, the Canadian Institute for Advanced Research, and the Miller Institute for Basic Research in Science (J.X.M.); and the NSF Earth System History program (P.U.C.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1111575/DC1
Figs. S1 to S3
References and Notes

28 February 2005; accepted 16 June 2005
Published online 23 June 2005;
10.1126/Science.1111575
Include this information when citing this paper.

Antagonistic Control of Disease Resistance Protein Stability in the Plant Immune System

Ben F. Holt III,^{1*} Youssef Belkhadir,^{1*} Jeffery L. Dangl^{1,2,3,4,†}

Pathogen recognition by the plant immune system is governed by structurally related, polymorphic products of disease resistance (*R*) genes. *RAR1* and/or *SGT1b* mediate the function of many *R* proteins. *RAR1* controls preactivation *R* protein accumulation by an unknown mechanism. We demonstrate that *Arabidopsis* *SGT1b* has two distinct, genetically separable functions in the plant immune system: *SGT1b* antagonizes *RAR1* to negatively regulate *R* protein accumulation before infection, and *SGT1b* has a *RAR1*-independent function that regulates programmed cell death during infection. The balanced activities of *RAR1* and *SGT1*, in concert with cytosolic HSP90, modulate preactivation *R* protein accumulation and signaling competence.

Specificity in the *Arabidopsis* immune system relies on ~125 polymorphic disease resistance (*R*) genes, many of which encode NB-LRR

proteins containing nucleotide binding sites and leucine-rich repeats. NB-LRR proteins “recognize” pathogen proteins that can con-

tribute to pathogen virulence in the absence of host recognition. When recognized by the plant, these are termed avirulence (*Avr*) proteins. Pathogens from various kingdoms trigger similar NB-LRR-mediated defense responses. Conserved plant proteins control NB-LRR signaling (1, 2). These include *RAR1*, *SGT1*, and cytosolic HSP90, each identified by recessive mutations and/or gene silencing in barley, *Arabidopsis*, potato, tobacco, and tomato (3–8).

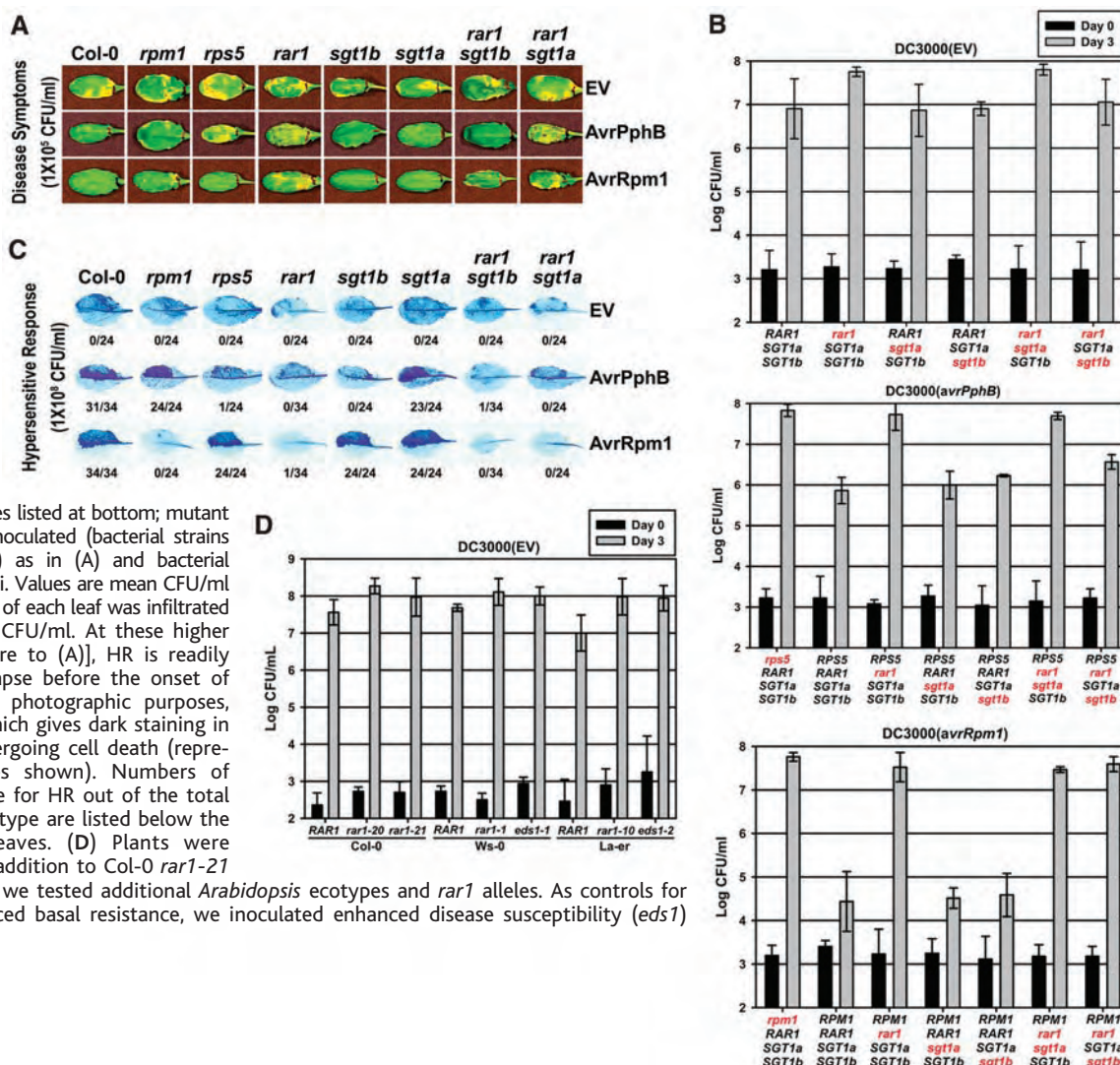
RAR1 plays a generic role in maintaining preactivation NB-LRR protein levels (9–11) (see below). However, *rar1* mutants suppress the resistance function of only a subset of

¹Department of Biology, ²Curriculum in Genetics, ³Department of Microbiology and Immunology, ⁴Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC 27599, USA.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. Department of Biology, 108 Coker Hall, University of North Carolina, CB# 3280, Chapel Hill, NC 27599–3280, USA. E-mail: dangl@email.unc.edu

Fig. 1. *SGT1b* antagonizes *RAR1* to control *RPS5*-mediated disease resistance. (A) *Pseudomonas syringae* pv. tomato (*Pto* DC3000) carrying empty vector (EV) or expressing *avrPphB* (to trigger *RPS5*) or *avrRpm1* (to trigger *RPM1*) was infiltrated into leaves at $\sim 1 \times 10^5$ colony forming units (CFU)/mL. Photos of disease symptoms were taken 5 days postinoculation (dpi). Plant lines, alternative alleles tested, extended protocols, and genotyping are described in (30). (B) Plants (genotypes listed at bottom; mutant loci in red) were hand inoculated (bacterial strains listed above each panel) as in (A) and bacterial growth was assessed 3 dpi. Values are mean CFU/mL ± 2 SE. (C) The upper half of each leaf was infiltrated as in (A) with 1×10^8 CFU/mL. At these higher inoculum levels [compare to (A)], HR is readily observed as tissue collapse before the onset of disease symptoms. For photographic purposes, we used trypan blue, which gives dark staining in regions of the leaf undergoing cell death (representative trypan leaves shown). Numbers of leaves scored as positive for HR out of the total examined for each genotype are listed below the trypan blue-stained leaves. (D) Plants were inoculated as in (A). In addition to *Col-0 rar1-21* [*rar1* allele used in (A)], we tested additional *Arabidopsis* ecotypes and *rar1* alleles. As controls for mutant lines with reduced basal resistance, we inoculated enhanced disease susceptibility (*eds1*) mutants.



NB-LRR proteins. A “threshold model” can explain the discrepancy between genetic requirements for RAR1 and its apparent biochemical function (11). Thus, *RAR1*–“independent” NB-LRR proteins accumulate to relatively high steady-state levels and remain above a threshold required for efficient defense activation even when destabilized in a *rar1* background. In contrast, *RAR1*–“dependent” NB-LRR proteins accumulate to relatively low levels that fall below a critical threshold in *rar1* mutants. Consistent with the semidominant nature of many *R*-mediated responses, the threshold model predicts that NB-LRR proteins are quantitative, response-limiting regulators. Cytosolic HSP90 is an additional determinant of steady-state NB-LRR protein accumulation (12). RAR1 likely collaborates with cytosolic HSP90 as a co-chaperone maintaining signal-competent NB-LRR proteins (13–16).

In yeast, SGT1 functions in kinetochore and SCF ubiquitin-ligase assembly (17–19). *Arabidopsis* has two SGT1 paralogs, *SGT1a* and *SGT1b* (78% amino acid identity), but only *sgt1b* mutations suppress NB-LRR function (7, 8, 20). RAR1, SGT1, and HSP90 interact in vivo, and RAR1 and SGT1 each interact with subunits of the COP9 signalosome, a likely proteasome lid complex (5, 14, 20). Further, SGT1 interacts with SCF ubiquitin ligase components, provoking speculation that SGT1 mediates the degradation of negative regulators of plant immune function (20). Concomitant losses of RAR1 and SGT1b additionally impair function of the *Arabidopsis* NB-LRR protein RPP5 (7), suggesting separable activities for these two genes. Accordingly, we define a RAR1-independent SGT1b function in programmed cell death. Unexpectedly, however, our data also demonstrate that SGT1b can negatively regulate NB-LRR protein accumulation, and that this activity is antagonized by both RAR1 and HSP90.

The *Arabidopsis* NB-LRR proteins RPM1, RPS2, and RPS5 confer resistance to *Pseudomonas syringae*. Each is impaired in *rar1* (10, 20, 21), but unaffected in *sgt1a* or *sgt1b* (7, 22) (Fig. 1, A and B). Unexpectedly, *RPS5* function, but not *RPM1* or *RPS2* function, was recovered in *rar1 sgt1b* (Fig. 1, A and B; *RPS2* data not shown). None of the *rar1* mutant phenotypes were recovered in *rar1 sgt1a*. Therefore, *SGT1b* mediates the loss of *RPS5* function in *rar1*, whereas *SGT1a* and *SGT1b* may act redundantly in this process for *RPM1* and *RPS2* (6).

NB-LRR activation often triggers a rapid localized programmed cell death, called the hypersensitive response (HR) (23). The HR likely limits the growth of biotrophic fungi and oomycetes (4, 21, 24, 25), although its role in resistance to bacterial pathogens is unclear. *RAR1* is required for *RPS5*-, *RPM1*-, and *RPS2*-mediated HR (10). Of these, only

the *RPS5*-mediated HR additionally required *SGT1b* (Fig. 1C; fig. S1A). Neither *RPS5*-, *RPM1*-, nor *RPS2*-dependent HR were restored in *rar1 sgt1b*. Using the oomycete parasite *Peronospora parasitica*, we extended these findings to two additional NB-LRR functions (*RPP4* and *RPP31*; fig. S1, B to E). Thus, *SGT1b* can control the HR in a *RAR1*-independent manner. Further, NB-LRR-mediated disease resistance and HR are genetically separable.

Notably, *rar1* mutations in different genetic backgrounds allowed enhanced growth of the virulent bacterial strain *P. syringae* (Pto) DC3000 (Fig. 1, B and D). These data demonstrate a role for *RAR1* in basal resistance, an ostensibly *R*-independent response that limits pathogen spread in susceptible plants (1). This *rar1* phenotype is also suppressed in *rar1 sgt1b*, but not *rar1 sgt1a* (Fig. 1D). Therefore, *SGT1b* also antagonizes *RAR1* in the control of basal resistance. Given that the only known function for *RAR1* is to promote NB-LRR protein accumulation, then NB-LRR

proteins also are very likely to function in basal resistance.

Requirements for *RAR1* and *SGT1b* have been defined for NB-LRR genes that confer resistance to different isolates of the oomycete parasite *Peronospora parasitica* (Pp) (table S1). *RPP8* was weakly impaired by *rar1*, as indicated by low levels of asexual parasite sporulation (Fig. 2, A and B). We bred isogenic plants hemizygous for an *RPP8* transgene (*RPP8*–) in each mutant background to determine whether the small phenotypic effect of *rar1* might depend on *RPP8* dosage. *RPP8*–/*rar1* plants exhibited increased susceptibility as compared to homozygous controls, supporting the threshold model (11). *RPP8*–/*rar1 sgt1b* plants were completely resistant, indicating that *SGT1b* mediates susceptibility in *RPP8*–/*rar1*. As with *RPP4*, *RPP31*, and *RPS5*, these data are inconsistent with the hypothesis that *RAR1* and *SGT1* act additively in all NB-LRR-mediated disease resistance responses.

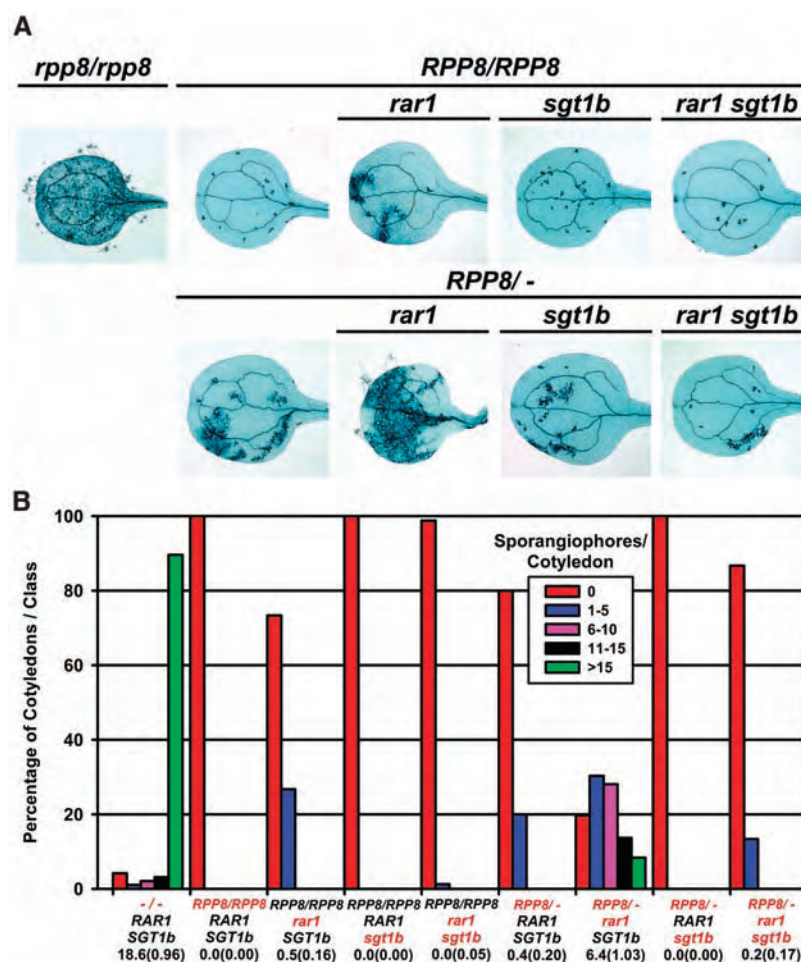
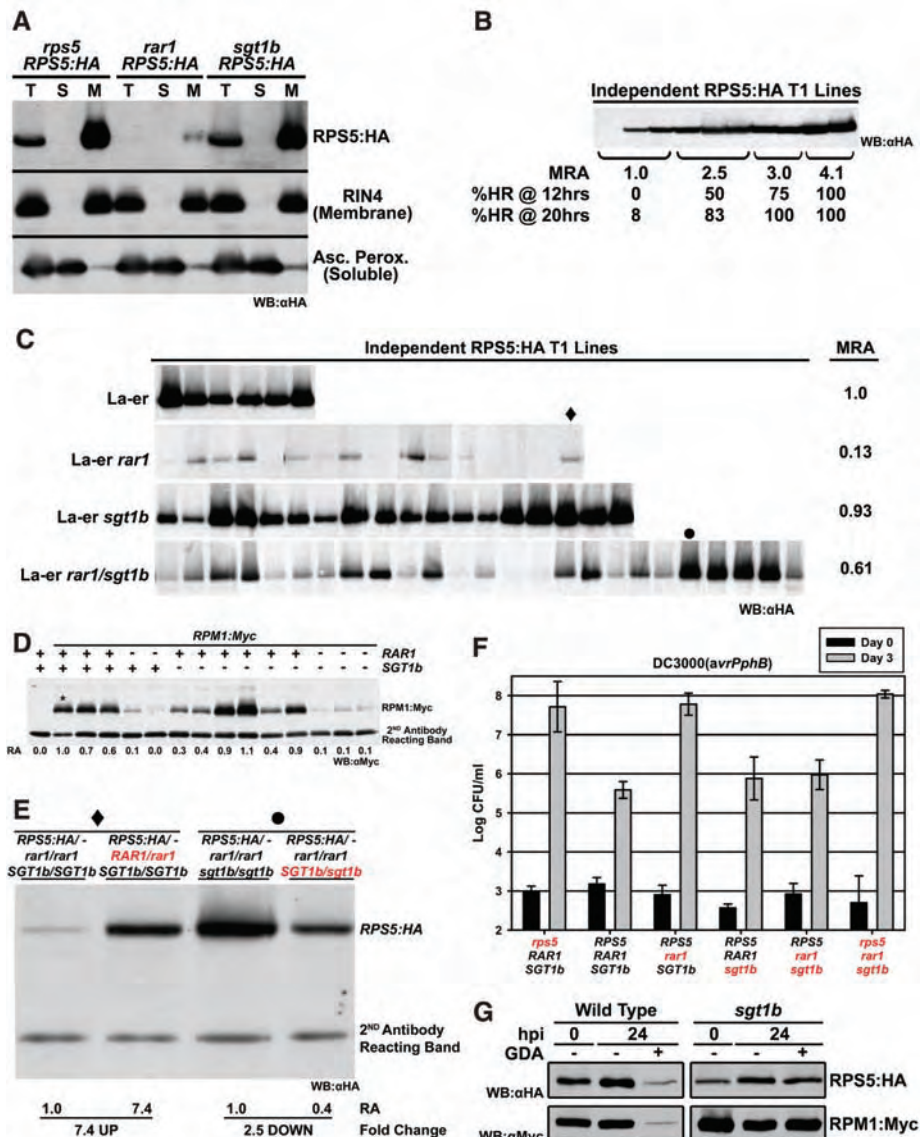


Fig. 2. SGT1b antagonism of RAR1 is generalizable to several NB-LRR resistance specificities. (A) Seven- to 10-day-old cotyledons of *rpp8* plants expressing a stable *RPP8* transgene were inoculated with the asexual spores of *Peronospora parasitica* (Pp) isolate Emco5 (40). Representative, trypan blue-stained leaves are shown to illustrate cell death and Pp structures (hyphae, asexual sporangioophores). (B) Asexual sporangioophores were quantified 7 dpi on at least 50 cotyledons for each of the indicated genetic backgrounds. The numbers below each tested genotype (key genotypes shown in red) represent mean sporangioophores/cotyledon (\pm 2 SE).

Unregulated NB-LRR expression can be lethal, suggesting that R protein accumulation must be fine tuned to provide rapid responses to infection while minimizing aberrant signaling. Dose dependence of *MLA1* (11) and *RPP8* (Fig.

We created genetic controls to confirm the antagonistic roles of RAR1 and SGT1b in RPS5 accumulation. A *rar1/rar1* transgenic parental line expressing low, but measurable RPS5:HA was used to generate *RPS5:HA*

Fig. 3. RAR1 and SGT1 act antagonistically to control RP55 protein accumulation. (A) Tissue samples for protein blot analysis were taken from independent, F_1 plants transformed with an HA epitope-tagged *RP55* transgene [*RP55:HA* (30)]. Protein was separated into total (T), soluble (S), and membrane (M) fractions (28). Ascorbate peroxidase and RIN4 antibodies were used as controls for the cytoplasmic and membrane fractions, respectively (41, 42). Equal loading for all protein samples in Fig. 3 was ensured by protein quantification before loading and Ponceau Red staining of nitrocellulose membranes after transfer. (B) Total protein extracts were isolated from 10 independent, F_1 Col-0 *rp55* mutants transformed with the *RP55:HA* transgene. Before protein blot analysis, four leaves per plant were visually scored for HR (as in Fig. 1C) at 12 and 20 hours (%HR@12 or 20 hrs). Mean relative *RP55:HA* protein accumulation (MRA) levels were quantified using ImageJ (version 1.31) (43). All values were transformed such that the weakest *RP55:HA*-expressing plants (first three lanes on blot) were equalized to MRA = 1.0. (C) La-er (*rp55*) ecotype plants and the *rar1*, *sgt1b*, and *rar1 sgt1b* mutants [also in La-er (30)] were transformed with the *RP55:HA* transgene. Individual, F_1 transformants were selected in each genetic background, and *RP55:HA* protein accumulation was visualized by protein blot. MRA values were transformed such that pooled values from the wild-type La-er ecotype was set to 1.0. Symbols above *rar1* and *rar1 sgt1b* lanes are explained in (E). (D) A stable RPM1:Myc transgenic line (28) was crossed to the *rar1 sgt1b* mutant. Indicated genotypes were selected by polymerase chain reaction from the F_2 population and examined by protein blot analysis as in (C). The lane designated with an asterisk (*) represents the parental RPM1:Myc line. (E) A La-er *RP55:HA/- rar1/rar1* transformant [male; (♦) in (C)] was crossed to either La-er *rar1/rar1* or La-er *RAR1/RAR1* (females in each cross). Similarly, a La-er *RP55:HA/- rar1/rar1 sgt1b/sgt1b* transformant [male; (●) in (C)] was crossed to either La-er *rar1/rar1 sgt1b/sgt1b* or La-er *rar1/rar1 SGT1b/SGT1b* (females). The resulting genotypes are shown above each lane. The first lane of each pair recapitulates the type, and the second represents altered gene *SGT1b* (red text). The secondary antibody reaction demonstrates equal loading. Relative accumulation formed such that the parental lane in each pair is 1.0. Stable, nonsegregating *rp55 rar1 sgt1b* triple-mutant



and tested for disease resistance as in Fig. 1B. (G) Leaves were infiltrated with either dimethyl sulfoxide (DMSO) alone or 10 μ M geldanamycin (GDA; A.G. Scientific, San Diego, CA) dissolved in DMSO (30). Samples were collected for protein blot analysis 24 hours after inoculation (similar results were seen at 18 hours). GDA did not alter RPS2:HA accumulation (data not shown) (30).

RAR1/rar1 and sibling control F_1 plants (Fig. 3E, first two columns). RPS5:HA accumulation was restored more than sevenfold in the *RAR1/rar1* heterozygote. Similarly, a *rar1 sgt1b* transgenic parent that accumulated high levels of RPS5:HA was used to generate *RPS5:HA rar1/rar1 SGT1b/sgt1b* and sibling control F_1 plants (Fig. 3E, third and fourth columns). The presence of a single copy of wild-type *SGT1b* resulted in 2.5 fold less RPS5:HA than in sibling controls. Importantly, disease resistance observed in *RPS5 rar1 sgt1b* (Fig. 1, A and B) was lost in an *rps5 rar1 sgt1b* triple mutant (Fig. 3F), demonstrating a direct link between restoration of RPS5 function and RPS5 protein levels. Collectively, these data demonstrate that RAR1 is a positive regulator, and SGT1b a negative regulator, of RPS5 accumulation. We envision that the recovery we observed for other NB-LRR functions in *rar1 sgt1b* (Fig. 2 and fig. S1, B to E) follows the same mechanism.

Reduction of cytosolic HSP90 function negatively affects steady-state accumulation of NB-LRR proteins (12, 14). We used the HSP90-specific inhibitor geldanamycin (GDA) (29) to examine RPS5:HA and RPM1:Myc protein accumulation in wild-type and *sgt1b* plants. GDA infiltration into wild-type leaves typically resulted in reduced RPS5:HA and RPM1:Myc protein accumulation, but did not eliminate disease resistance function (Fig. 3G) (30). GDA did not affect accumulation of either NB-LRR protein in *sgt1b*. Thus, elimination of RAR1 or inhibition of HSP90 activity is sufficient to lower NB-LRR protein accumulation through an unknown mechanism. In both cases, SGT1b can mediate this outcome. Notably, RPM1:Myc destabilization mediated by GDA is SGT1b dependent, whereas its destabilization in *rar1* is not. This contrasts with RPS5:HA, suggesting that antagonism between RAR1-HSP90 and SGT1b is fine tuned for different NB-LRR proteins.

Our findings challenge suggestions of signaling functions for RAR1 and SGT1b in NB-LRR-mediated disease resistance. Restoration of *RPS5*-, *RPP4*-, *RPP8*-, and *RPP31*-mediated functions in *rar1 sgt1b* prove that RAR1 and SGT1b are not required for disease resistance signaling per se. Additionally, we show that SGT1b has a RAR1-independent function as a positive regulator of *RPP4*-, *RPP31*-, and *RPS5*-mediated HR. A general role for SGT1b in HR is now well established (6, 31), and we speculate that an efficient HR requires SGT1b-dependent elimination of an unidentified negative regulator. This SGT1b function would be particularly relevant in cases where HR plays a key role in limiting pathogen spread, explaining why some NB-LRR proteins exhibit additive requirements for RAR1 and SGT1b. In such cases, the lack of NB-LRR accumulation in *rar1 sgt1b* coupled to an inefficient HR would result in enhanced pathogen growth.

RAR1 and HSP90 are positive regulators of NB-LRR protein steady-state accumulation [(9–12, 14) and this work]. As such, RAR1 and HSP90 may determine whether NB-LRR proteins are functional in disease resistance or marked for degradation. Cytosolic HSP90 transiently binds nonnative “client” proteins to assist in proper folding (32, 33). Active folding of HSP90 client proteins is regulated by cycles of adenosine 5'-triphosphate (ATP) binding and hydrolysis that are, in turn, modulated by co-chaperones. In addition to modulating ATP hydrolysis, co-chaperones also guide HSP90 client specificity. Therefore, HSP90 apparently processes and/or maintains NB-LRR proteins to a signal-competent conformational state, with RAR1 acting as a co-chaperone.

Yeast SGT1 transiently links HSP90 to the inner kinetochore complex (CBF3), balancing CBF3 assembly and turnover (34). Specific mutations that “trap” SGT1 in CBF3 complexes result in reduced CBF3 accumulation. This is consistent with our finding that elimination of SGT1b can reduce NB-LRR turnover. We speculate that RAR1 defines a regulatory checkpoint protecting HSP90-associated NB-LRR proteins from SGT1b-mediated degradation. In *rar1* mutants, this degradation pathway becomes the default, perhaps through direct interaction of HSP90-associated NB-LRR proteins with an SCF-bound SGT1 (11, 35, 36).

Coupling of folding and degradation fates has previously been demonstrated for the HSP90 clients glucocorticoid hormone receptor (GR) and cystic fibrosis transmembrane conductance regulator (CFTR) (37, 38). GR or CFTR, in complex with HSP70/HSP90, are degraded when these complexes associate with CHIP (carboxy-terminus of HSP70 interacting protein), a member of the U-box family of ubiquitin ligases. Mutations in CHIP that eliminate ubiquitin ligase function dominantly interfere with ubiquitination and subsequent GR/CFTR degradation. Like SGT1, CHIP has several tetratricopeptide repeats (TPRs) that are required for HSP70/HSP90 association (15, 19, 37). Therefore, like CHIP, SGT1-SCF complexes might couple NB-LRR proteins to the cellular degradation machinery (39). It remains unclear whether changes in NB-LRR accumulation are due to proteasome-dependent degradation or an alternative protein turnover mechanism such as endocytosis. Nevertheless, we anticipate that our genetic results will inform subsequent biochemical experiments.

References and Notes

1. J. L. Dangl, J. D. G. Jones, *Nature* **411**, 826 (2001).
2. Y. Belkhadir, R. Subramaniam, J. L. Dangl, *Curr. Opin. Plant Biol.* **7**, 391 (2004).
3. Y. Liu, M. Schiff, R. Marathe, S. P. Dinesh-Kumar, *Plant J.* **30**, 415 (2002).
4. K. Shirasu et al., *Cell* **99**, 355 (1999).
5. Y. Liu, M. Schiff, G. Serino, X. W. Deng, S. P. Dinesh-Kumar, *Plant Cell* **14**, 1483 (2002).
6. J. R. Peart et al., *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10865 (2002).

7. M. J. Austin et al., *Science* **295**, 2077 (2002).
8. M. Tör et al., *Plant Cell* **14**, 993 (2002).
9. Y. Belkhadir, Z. Nimchuk, D. A. Hubert, D. Mackey, J. L. Dangl, *Plant Cell* **16**, 2822 (2004).
10. P. Torner et al., *Plant Cell* **14**, 1005 (2002).
11. S. Bieri et al., *Plant Cell* **16**, 3480 (2004).
12. R. Lu et al., *EMBO J.* **22**, 5690 (2003).
13. Y. Liu, T. Burch-Smith, M. Schiff, S. Feng, S. P. Dinesh-Kumar, *J. Biol. Chem.* **279**, 2101 (2004).
14. D. A. Hubert et al., *EMBO J.* **22**, 5679 (2003).
15. A. Takahashi, C. Casais, K. Ichimura, K. Shirasu, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11777 (2003).
16. K. Shirasu, P. Schulze-Lefert, *Trends Plant Sci.* **8**, 252 (2003).
17. K. Kitagawa, D. Skowyrza, S. J. Elledge, J. W. Harper, P. Hieter, *Mol. Cell* **4**, 21 (1999).
18. P. Steensgaard et al., *EMBO Rep.* **5**, 626 (2004).
19. P. K. Bansal, R. Abdulle, K. Kitagawa, *Mol. Cell. Biol.* **24**, 8069 (2004).
20. C. Azevedo et al., *Science* **295**, 2073 (2002).
21. P. R. Muskett et al., *Plant Cell* **14**, 979 (2002).
22. R. F. Warren, P. M. Merritt, E. Holub, R. W. Innes, *Genetics* **152**, 401 (1999).
23. J. L. Dangl, R. A. Dietrich, M. H. Richberg, *Plant Cell* **8**, 1793 (1996).
24. L. Belbahri et al., *Plant J.* **28**, 419 (2001).
25. A. Freialdenhoven et al., *Plant Cell* **6**, 983 (1994).
26. M. T. Simonich, R. W. Innes, *Mol. Plant Microbe Interact.* **8**, 637 (1995).
27. M. J. Axtell, B. J. Staskawicz, *Cell* **112**, 369 (2003).
28. D. C. Boyes, J. Nam, J. L. Dangl, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15849 (1998).
29. S. M. Roe et al., *J. Med. Chem.* **42**, 260 (1999).
30. Materials and methods are available as supporting material on Science Online.
31. Y. Zhang, S. Dorey, M. Swiderski, J. D. Jones, *Plant J.* **40**, 213 (2004).
32. J. C. Young, I. Moarefi, F. U. Hartl, *J. Cell Biol.* **154**, 267 (2001).
33. T. A. Sangster, C. Queitsch, *Curr. Opin. Plant Biol.* **8**, 86 (2005).
34. L. B. Lingelbach, K. B. Kaplan, *Mol. Cell. Biol.* **24**, 8938 (2004).
35. C. Dubacq, R. Guerois, R. Courbeyrette, K. Kitagawa, C. Mann, *Eukaryot. Cell* **1**, 568 (2002).
36. Yeast SGT1 interacts with the LRR of Cyp1p/Cdc35p adenyl cyclase, and barley SGT1 interacts with the LRR of barley NB-LRR protein MLA1. However, full-length MLA1 does not interact with SGT1, consistent with the suggestion that NB-LRR conformation is regulated by intramolecular interactions that differ between pre- and postinfection states. We found that the LRR domains of RPM1, RPS2, and RPS5 did not interact with either SGT1a or SGT1b in yeast two-hybrid experiments, although we were able to reproduce the MLA1-SGT1 interaction. We also found no obvious proteasome function in RPS5 degradation (30).
37. P. Connell et al., *Nat. Cell Biol.* **3**, 93 (2001).
38. G. C. Meacham, C. Patterson, W. Zhang, J. M. Younger, D. M. Cyr, *Nat. Cell Biol.* **3**, 100 (2001).
39. J. Höfelfeld, D. M. Cyr, C. Patterson, *EMBO Rep.* **2**, 885 (2001).
40. J. M. McDowell et al., *Plant Cell* **10**, 1861 (1998).
41. H. M. Jespersen, I. V. Kjaersgaard, L. Ostergaard, K. G. Welinder, *Biochem. J.* **326**, 305 (1997).
42. D. Mackey, B. F. Holt III, A. Wiig, J. L. Dangl, *Cell* **108**, 743 (2002).
43. <http://rsb.info.nih.gov/ij/>
44. Supported by NSF Arabidopsis 2010 grant (IBN-0114795). We thank D. Hubert, J. Chang, J. McDowell, E. Holub, and J. Jones for critical evaluations of the manuscript. We also thank P. Muskett and J. Parker for providing *La-er rar1-10* and *sgt1b-1* seeds and W. Gray for *sgt1b^{enh3}* seeds. L. C. Tran provided technical support.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1109977/DC1
Materials and Methods
Figs. S1 and S2
Table S1
References and Notes

19 January 2005; accepted 9 June 2005
Published online 23 June 2005;
10.1126/science.1109977
Include this information when citing this paper.

Stem Cell Depletion Through Epidermal Deletion of Rac1

Salvador Aznar Benitah,¹ Michaela Frye,¹ Michael Glogauer,²
Fiona M. Watt^{1*}

Mammalian epidermis is maintained by self-renewal of stem cells, but the underlying mechanisms are unknown. Deletion of Rac1, a Rho guanosine triphosphatase, in adult mouse epidermis stimulated stem cells to divide and undergo terminal differentiation, leading to failure to maintain the interfollicular epidermis, hair follicles, and sebaceous glands. Rac1 exerts its effects in the epidermis by negatively regulating c-Myc through p21-activated kinase 2 (PAK2) phosphorylation. We conclude that a pleiotropic regulator of cell adhesion and the cytoskeleton plays a critical role in controlling exit from the stem cell niche and propose that Rac and Myc represent a global stem cell regulatory axis.

Mammalian epidermis is maintained by proliferation of stem cells and differentiation of their progeny along the lineages of the interfollicular epidermis (IFE), sebaceous gland (SG), and hair follicle (HF) (1). Although studies of cultured keratinocytes have established the importance of extracellular matrix adhesion in regulating the onset of terminal differentiation, targeted deletion of the major keratinocyte integrins *in vivo* does not give the stem cell depletion phenotypes anticipated (2). Epidermal stem cell depletion does occur on activation of c-Myc (3–5), which not only causes reduced integrin expression but also decreased expression of actin cytoskeleton components (6). This suggests that maintenance of the stem cell compartment depends not only on integrin engagement but also on pleiotropic downstream effectors that affect the cytoskeleton, such as Rac1. We therefore investigated the consequences of epidermal deletion of Rac1 (7, 8), by applying 4-hydroxy-tamoxifen (4OHT) to the skin of mice expressing floxed Rac1 alleles (9) together with CreER under the control of the keratin 14 promoter (10).

In wild-type mouse epidermis, Rac1 protein (Fig. 1A and fig. S1A) and mRNA (Fig. 1B) were highly expressed in the basal, undifferentiated layer of the IFE. As expected, 4OHT treatment of transgenic mice led to the loss of Rac1 expression (Fig. 1C and fig. S2A). In wild-type HFs, staining for Rac1 was most intense in the bulge, which constitutes a stem cell reservoir (11, 12), and at the base of the follicle, the bulb (Fig. 1D and fig. S1A). Rac1 expression in the bulge did not change during the hair cycle, but expression in the bulb expanded as the bulb enlarged in growing

follicles (13). Rac1 expression was elevated in papillomas and squamous cell carcinomas (fig. S1, B to D), which are believed to arise from epidermal stem cells (1).

In human IFE, there was strong expression of Rac1 in the basal layer, and colocalization with the hemidesmosomal integrin $\alpha 6 \beta 4$ at the ventral plasma membrane (Fig. 1, E and F). Rac1 expression was particularly high in clusters of basal cells enriched in stem cells (14)

(fig. S1, E and F) and was elevated in squamous cell carcinomas (fig. S1, G and H).

When cultured human epidermal cells were transduced with Rac1 RNA interference (RNAi), clonal growth was greatly reduced (Fig. 1G), and expression of terminal differentiation markers such as transglutaminase 1 increased (fig. S3H). Conversely, transduction with active Rac1 (Rac1QL) increased the proportion of cells capable of clonal growth (Fig. 1G).

The dorsal skin of K14CreER/floxed Rac1 mice treated with 4OHT showed three distinct phenotypes, designated early, middle, and late (Fig. 1, H to K, and figs. S2, B to D and S3, A to D). After 3 to 5 days (early), there was thickening of the IFE with increased numbers of living and cornified cell layers (Fig. 1, H and I), and the infundibulum, at the junction between the IFE and HF, was expanded (fig. S3B). After 7 to 9 days (middle), there was disorganization and decreased cellularity of the IFE basal layer, together with cell enlargement (Fig. 1J). SGs were also enlarged and disorganized (fig. S3C). After 11 to 15 days, the late phenotype developed: partial or complete loss of viable IFE cell layers (Fig. 1K), diminution of the HF bulb, and degeneration of the infundibulum into cysts (fig. S3D). Rac1 deletion also caused pronounced defects in the HF growth cycle (fig. S2E). In the early stage of

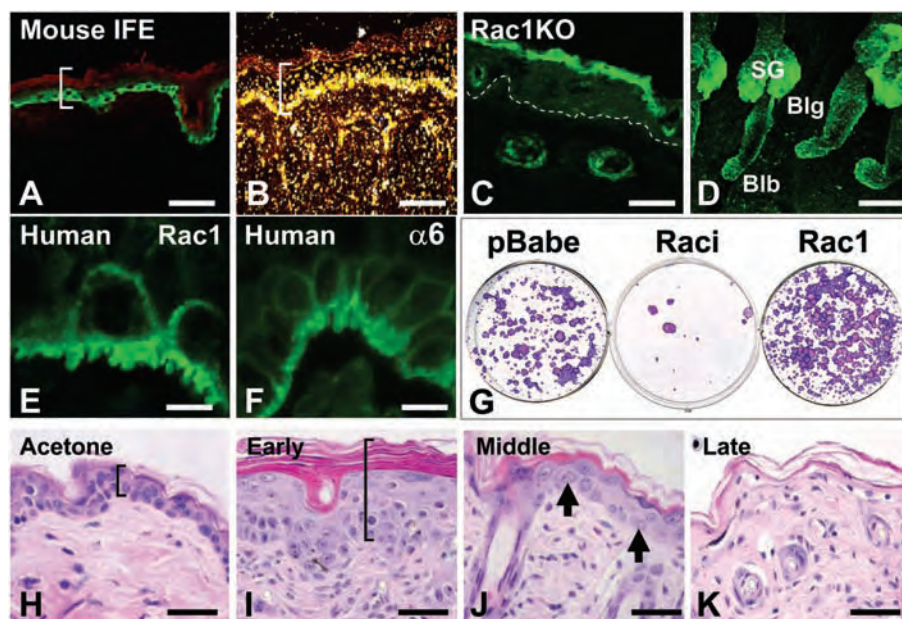


Fig. 1. Epidermal Rac1 expression and effects of Rac1 deletion. (A) Rac1 expression in wild-type mouse back skin (green fluorescence, Rac1; red fluorescence, involucrin, a terminal differentiation marker; the bracket denotes IFE). (B) Rac1 mRNA in tongue epithelium (the bracket denotes IFE). (C) Absence of Rac1 protein in K14CreER/floxed Rac1 (Rac1KO) dorsal epidermis treated with 4OHT for 7 days (the dotted line indicates the basement membrane). (D) Rac1 protein in wild-type mouse tail epidermis (wholemount; SG, sebaceous gland; Blg, bulge; Blb, bulb; SG staining is nonspecific). (E and F) Human IFE basal layer with (E) Rac1 expression and (F) hemidesmosomal integrin $\alpha 6 \beta 4$ ($\alpha 6$) expression. (G) Clonal growth of primary human keratinocytes transduced with empty vector (pBabe), Rac1 RNAi (Raci), or Rac1QL (Rac1). (H to K) Hematoxylin and eosin-stained sections of K14CreER/floxed Rac1 back skin treated with acetone (control) or 4OHT to induce the early, middle, and late phenotypes. Brackets denote IFE; arrows, cells in the IFE basal layer. Scale bars, [(A) to (C) and (H) to (K)], 100 μ m; (D) 2 μ m; [(E) and (F)], 1 μ m.

¹Keratinocyte Laboratory, Cancer Research UK London Research Institute, 44 Lincoln's Inn Fields, London WC2A 3PX, UK. ²Faculty of Dentistry, University of Toronto, College Street, Toronto M5S 1A8, Canada.

*To whom correspondence should be addressed. E-mail: fiona.watt@cancer.org.uk

Rac1 deletion, the size of hemidesmosomes decreased (fig. S3, E and F); subsequently, hemidesmosomes were reduced in number and rudimentary in structure (fig. S3G).

The phenotype of Rac1 deletion in vivo and in culture suggested that exit from the stem cell compartment was induced. If so, a transient increase in proliferation would be expected, as cells enter the transit-amplifying compartment, followed by a decrease as transit-amplifying divisions are exhausted and cells undergo terminal differentiation (1, 2). At 3 to 5 days after 4OHT treatment, there was increased proliferation (Fig. 2, A and E), whereas from day 7 onwards there were fewer Ki67-positive cells than in control epidermis (Fig. 2I and fig. S4A). The number of suprabasal layers expressing the IFE differentiation marker, keratin 10 (15), increased in the early phase of Rac1 deletion (Fig. 2, B and F); at later times, keratin 10 expression extended into the basal layer (Fig. 2J and fig. S4B). Consistent with the loss of hemidesmosomes (fig. S3G), $\alpha 6 \beta 4$ integrin expression was progressively reduced (Fig. 2, C, G, and K, and fig. S4C). Whole mounts of tail epidermis revealed an early increase in terminally differentiated sebocytes (Fig. 2, D and H), reminiscent of the consequences of activating Myc (3, 6, 15), followed by progressive sebocyte loss (Fig. 2L and fig. S4D). Rac1 deletion led to thickening of the infundibulum (Fig. 2L, arrow) and loss of the bulb, confirmed by reduced CDP (CCAAT displacement protein) expression (15) (fig. S4, E and F). Although Rac1 is a potent pro-survival signal (7), there was no increase in apoptosis in Rac1 null epidermis (fig. S4, G and H).

To confirm that Rac1 deletion led to depletion of the stem cell compartment, we examined three HF bulge markers: K15, CD34, and high $\alpha 6 \beta 4$ integrin expression (11, 12, 15). All showed a substantial reduction after prolonged 4OHT treatment (Fig. 3, A to F, and fig. S5, A to D). By 9 days, the number of CD34+ cells expressing high levels of $\alpha 6$ decreased from 8% to 4% (Fig. 3, A and B), and cell surface levels of each protein were reduced (Fig. 3, C and D). $\beta 1$ integrins, which are uniformly expressed along the HF (15), were unaffected by Rac1 loss (fig. S5, E and F), providing an explanation for why epidermal detachment from the basement membrane did not occur.

A subpopulation of HF bulge stem cells can be identified as bromodeoxyuridine label-retaining cells (LRCs) (16). During the early response to Rac1 deletion, LRCs divided (Fig. 3, G and H) and the LRC zone expanded, extending from the infundibulum to the bulb (Fig. 3, I and J, brackets). The proportion of LRCs that were in the S (2.6% versus 0.7%) or G2+M (5.5% versus 2.9%) phase of the cell cycle increased relative to control epidermis (Fig. 3, G and H). Continued proliferation of LRCs (Fig. 3K) eventually resulted in loss of

the label, correlating with the appearance of infundibulum cysts (Fig. 3L, arrow).

Additional evidence for stem cell depletion came from the clonal growth of epidermal cells in culture. Treatment of cultured K14CreER/floxed Rac1 keratinocytes with 4OHT completely inhibited clonal growth (fig. S5G), and colony formation was rescued by Rac1QL (Fig. 1G and fig. S5G).

Rac1 exerts its biological effects through specific downstream effectors, such as PAK2 (8), which negatively regulates c-Myc (17). Because Myc activation causes epidermal stem cell depletion in vivo (3, 4) and in culture (18), we investigated whether Rac1 deletion affected Myc expression. During the early phase of Rac1 deletion, c-Myc expression increased (Fig. 4, A to C, and fig. S6A); conversely, when Myc was activated in K14MycER transgenic mice (3), expression of Rac1 was down-regulated (Fig. 4D and fig. S6B). Overexpression of activated Rac1 in human keratinocytes blocked Myc-induced terminal differentiation (Fig. 4E). Activated Rac1 prevented the decrease in clonal growth (Fig. 4F) and $\alpha 6$ integrin expression (fig. S6D) that occurs on Myc activation (6, 18).

Phosphorylated PAK2 colocalized with Rac1 in the basal layer of human IFE (fig. S6, E and F), and PAK2 bound and phosphorylated c-Myc in keratinocytes (fig. S6C). When Rac1QL was introduced into cultured hu-

man keratinocytes, phosphorylation of PAK2 increased (fig. S6G). To investigate whether Rac1 blocked the effects of Myc through PAK2, we introduced two 4OHT-inducible Myc mutants into cultured human keratinocytes: MycAER (T358A/S373A/T400A), which cannot be phosphorylated by PAK2, and MycDER (T358D/S373D/T400D), which mimics constitutive phosphorylation of PAK2 sites (17). The effects of MycAER on epidermal differentiation (Fig. 4G) and $\alpha 6$ integrin expression (Fig. 4I) were similar to the effects of wild-type MycER (Fig. 4, G and H). In contrast, MycDER did not stimulate differentiation (Fig. 4G) and increased $\alpha 6$ integrin expression (Fig. 4J).

Activation of MycER increases the proportion of transit-amplifying cells, which give rise to abortive clones consisting of large, terminally differentiated keratinocytes (18) (fig. S6, H and I). Clones of cultured human keratinocytes expressing MycAER resembled MycER clones (fig. S6, I and J). In contrast, MycDER clones (fig. S5K) resembled clones expressing Rac1QL (fig. S6L), with a high proportion of small, undifferentiated keratinocytes. Rac1QL rescued MycER clones from differentiation (fig. S6M), and the combination of Rac1QL and MycDER also gave rise to undifferentiated clones (fig. S6O). However, MycAER blocked the action of constitutively active Rac1 (fig. S6N), supporting the conclusion that the antagonistic effect of Rac1 on Myc involves PAK2.

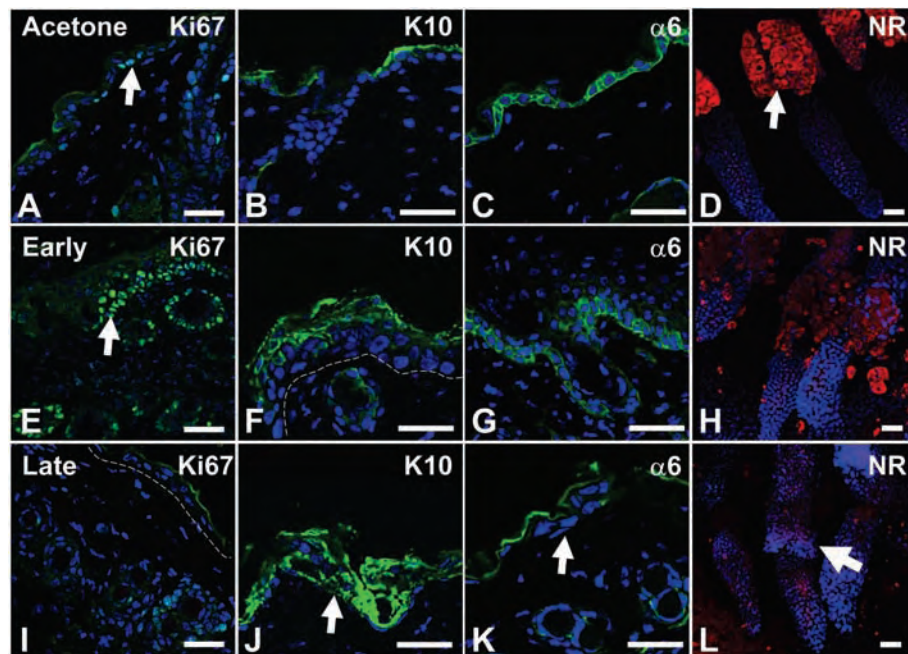


Fig. 2. Deletion of Rac1 triggers transient proliferation followed by terminal differentiation. K14CreER/floxed Rac1 mice were treated with (A to D) acetone (control) or with 4OHT to induce (E to H) early or (I to L) late phenotypes. Immunostaining is shown for Ki67, keratin 10 (K10), and $\alpha 6$ integrin in sections of dorsal epidermis. The arrows in (A) and (E) denote Ki67-positive cells; the arrow in (D) denotes the SG; the arrow in (J) denotes K10-positive basal cells; and the arrow in (L) denotes the infundibulum cyst. The dotted lines indicated the basement membrane. Differentiated sebocytes were detected with Nile Red (NR) in whole mounts of tail epidermis. Blue fluorescence, 4',6-diamidino-2-phenylindole (DAPI) nuclear counterstain. In (H), sebocytes detached from the SG during processing. Scale bars, 100 μ m.

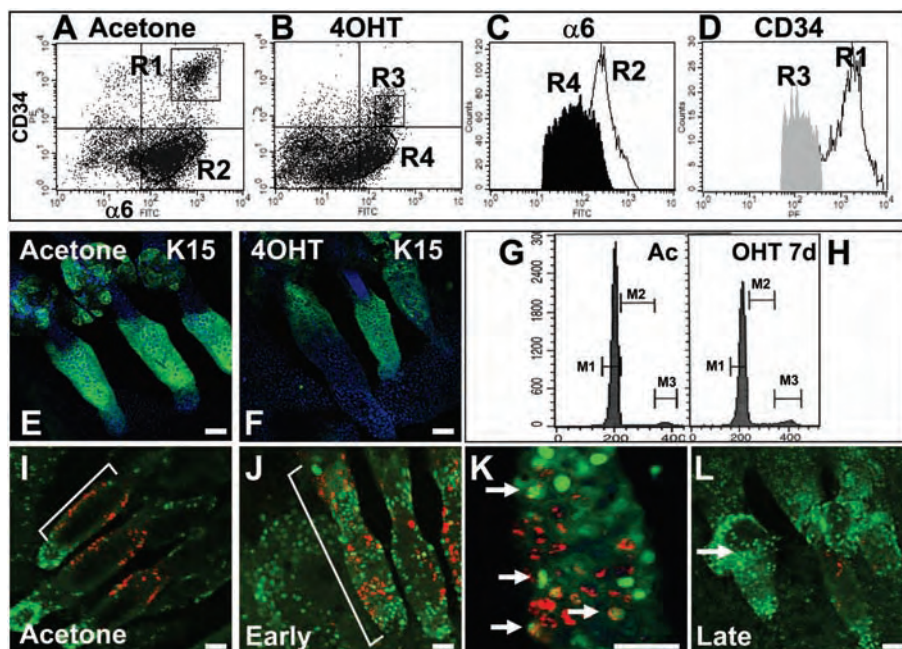


Fig. 3. Depletion of epidermal stem cells. (A to D) Flow cytometry of $\alpha 6$ integrin and CD34 double-labeled keratinocytes from K14CreER/floxed Rac1 back skin treated with acetone or 4OHT (for 7 days). FITC, fluorescein isothiocyanate; PE, phycoerythrin. The cell populations denoted by R1 through R4 in (A) and (B) are also shown in (C) and (D). (E and F) Keratin 15 immunofluorescence of tail epidermal whole mounts from K14CreER/floxed Rac1 mice treated with acetone or 4OHT (late phenotype). (G and H) Propidium-labeled LRCs. Ac, acetone; 7d, 7 days. Horizontal lines designated M1, M2, and M3 denote cells with G1, S, and G2+M DNA content, respectively. (I to L) Tail epidermal whole mounts from K14CreER/floxed Rac1 mice treated with acetone or 4OHT [(I) and (K), early phenotype; (L) late phenotype]. LRCs are shown in green, and Ki67 in red. The LRC zone is demarcated by brackets in (I) and (J); arrows show Ki67-positive LRCs in (K) and the infundibulum cyst in (L). Scale bars, 100 μ m.

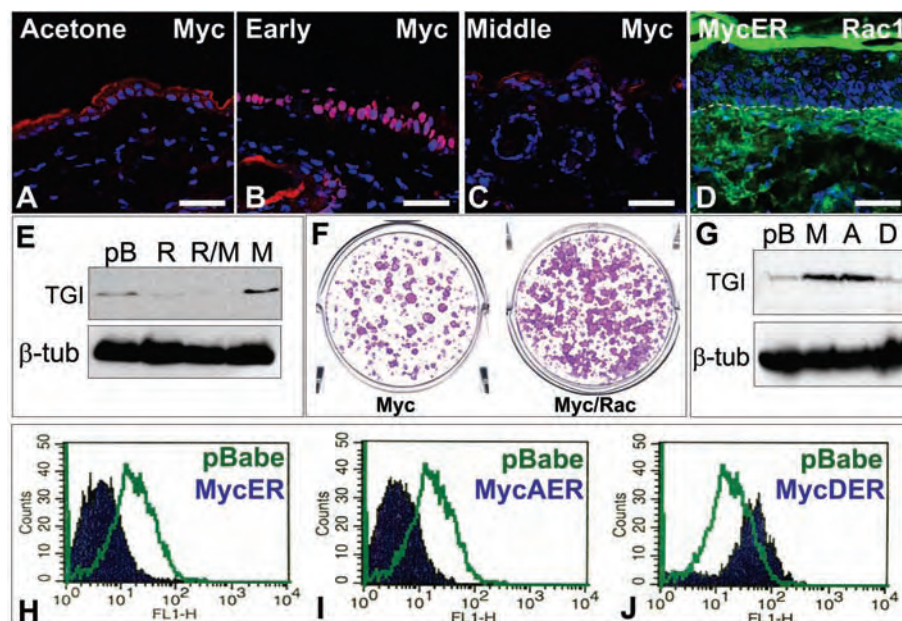


Fig. 4. Rac1 inhibits c-Myc through PAK2. (A to C) K14CreER/floxed Rac1 or (D) K14MycER epidermis treated with acetone or 4OHT and stained with antibody to c-Myc (red), Rac1 (green), and DAPI (blue). The dotted line indicates the basement membrane. (E and F) Rac1 rescues the effect of c-Myc on (E) human keratinocyte terminal differentiation and (F) clonal growth. Cells were transduced with empty vector (pB), Rac1QL (R and Rac), and MycER (M and Myc), alone or together and treated with 4OHT. (E) A Western blot probed for transglutaminase 1 (TGI) or β -tubulin (β -tub). (G to J) Effects of MycER (M), MycAER (A), and MycDER (D) on (G) transglutaminase levels and [(H) to (J)] surface $\alpha 6$ integrin expression. FL1-H, fluorescence attributable to binding of antibody to $\alpha 6$. Scale bar, 100 μ m.

We conclude that deletion of Rac1 from adult mouse epidermis leads to rapid depletion of stem cells. Rac1 is a pleiotropic regulator of many cellular processes, including integrin and growth factor signaling and cell-cell adhesion (7, 8). Nevertheless, one key mechanism by which Rac1 maintains epidermal stem cells is by negatively regulating Myc through PAK2 phosphorylation. Although Myc is a proto-oncogene, it promotes differentiation of epidermal and hemopoietic lineages, disrupting adhesive interactions between stem cells and their niche (5). In these tissues, either Rac deletion (19, 20) or Myc activation (5) depletes the stem cell compartment. Conversely, in intestinal epithelium, Myc promotes self-renewal (5) and Rac1 stimulates differentiation (21). Thus, although their precise roles are undoubtedly dependent on cellular context, Rac and Myc appear to represent a global stem cell regulatory axis.

The consequences of Rac1 deletion in the epidermis demonstrate that cell adhesion and the cytoskeleton regulate not only tissue organization but also differentiation. Furthermore, they suggest that, in addition to promoting invasion (7, 8), increased expression of Rac1 in epithelial tumors may stimulate expansion of the stem cell compartment and inhibit differentiation.

References and Notes

- D. M. Owens, F. M. Watt, *Nat. Rev. Cancer* **3**, 444 (2003).
- F. M. Watt, *EMBO J.* **21**, 3919 (2002).
- I. Arnold, F. M. Watt, *Curr. Biol.* **11**, 558 (2001).
- R. L. Waikel, Y. Kawachi, P. A. Waikel, X. J. Wang, D. R. Roop, *Nat. Genet.* **28**, 165 (2001).
- M. J. Murphy, A. Wilson, A. Trumpp, *Trends Cell Biol.* **15**, 128 (2005).
- M. Frye, C. Gardner, E. R. Li, I. Arnold, F. M. Watt, *Development* **130**, 2793 (2003).
- S. Etienne-Manneville, A. Hall, *Nature* **420**, 629 (2002).
- S. A. Benitah, P. F. Valeron, L. van Aelst, C. J. Marshall, J. C. Lacal, *Biochim. Biophys. Acta* **1705**, 121 (2004).
- M. Glogauer et al., *J. Immunol.* **170**, 5652 (2003).
- K. U. Hong, S. D. Reynolds, S. Watkins, E. Fuchs, B. R. Stripp, *Am. J. Pathol.* **164**, 577 (2004).
- T. Tumber et al., *Science* **303**, 359 (2004).
- R. J. Morris et al., *Nat. Biotechnol.* **22**, 411 (2004).
- S. Aznar Benitah, F. M. Watt, unpublished data.
- J. Legg, U. B. Jensen, S. Broad, I. Leigh, F. M. Watt, *Development* **130**, 6049 (2003).
- K. M. Braun et al., *Development* **130**, 5241 (2003).
- R. Morris, C. S. Potten, *J. Invest. Dermatol.* **112**, 470 (1999).
- Z. Huang, J. A. Traugh, J. M. Bishop, *Mol. Cell. Biol.* **24**, 1582 (2004).
- A. Gandarillas, F. M. Watt, *Genes Dev.* **11**, 2869 (1997).
- Y. Gu et al., *Science* **302**, 445 (2003).
- M. J. Walmsley et al., *Science* **302**, 459 (2003).
- T. S. Stappenbeck, K. I. Gordon, *Development* **127**, 2629 (2000).
- This work was funded by Cancer Research UK, by the European Union (EU) EuroStemCell network, and by European Molecular Biology Organization and EU Marie Curie fellowships (S.A.B.). We thank everyone who provided advice and technical support.

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5736/933/DC1

Materials and Methods

Figs. S1 to S5

14 April 2005; accepted 13 June 2005

10.1126/science.1113579

Protein Structures Forming the Shell of Primitive Bacterial Organelles

Cheryl A. Kerfeld,^{1,2,4} Michael R. Sawaya,^{1,3,4} Shiho Tanaka,¹
Chau V. Nguyen,¹ Martin Phillips,^{1,3} Morgan Beeby,^{1,3}
Todd O. Yeates^{1,3,4*}

Bacterial microcompartments are primitive organelles composed entirely of protein subunits. Genomic sequence databases reveal the widespread occurrence of microcompartments across diverse microbes. The prototypical bacterial microcompartment is the carboxysome, a protein shell for sequestering carbon fixation reactions. We report three-dimensional crystal structures of multiple carboxysome shell proteins, revealing a hexameric unit as the basic microcompartment building block and showing how these hexamers assemble to form flat facets of the polyhedral shell. The structures suggest how molecular transport across the shell may be controlled and how structural variations might govern the assembly and architecture of these subcellular compartments.

Intracellular polyhedral bodies (Fig. 1, A and B) were first observed in bacteria by electron microscopy more than 40 years ago (1–3). These large protein shells, generally ranging in size from 100 to 200 nm, were initially thought to be viral particles (4), but the first such structures isolated from chemoautotrophic bacteria (5) were filled with the enzyme ribulose biphosphate carboxylase oxygenase (RuBisCO) and were therefore named carboxysomes. In organisms where they occur, carboxysomes compartmentalize most, if not all, of the cellular RuBisCO. Subsequent biochemical characterization demonstrated that carboxysomes also have associated carbonic anhydrase activity, providing localized conversion of bicarbonate to CO₂ (6–8), the substrate for RuBisCO. Carbon fixation may be optimized by providing a microenvironment rich in CO₂ within a shell that could be less permeable to RuBisCO's potent competitive inhibitor, molecular oxygen (9–11).

Comparative biochemical and genomic analysis of carboxysomes has revealed that several small (~10 kD) highly similar homologous proteins (Fig. 2) are the key components of the carboxysome shell (11, 12). Evolutionarily related proteins have been identified in several obligately heterotrophic enteric bacteria (which do not fix CO₂), encoded within operons associated with oxygen-sensitive enzymatic processes (11, 13–15). In

those organisms, ultrastructural studies confirm the presence of carboxysome-like inclusions under growth conditions that induce those operons. Collectively, the shell proteins of bacterial microcompartments contain a conserved sequence referred to as the bacterial microcompartment (BMC) domain. Querying the sequence databases for similarity to the BMC domain reveals that the potential to form proteinaceous compartments is widespread among the bacteria (fig. S1). Typically, several genes coding for BMC domain-containing proteins cluster with genes coding for putative enzymes or proteins of unknown function; presumably, they form specialized compartments

for as-yet uncharacterized metabolic processes. Thus, the microcompartment structure typified by the carboxysome can be viewed as a form of primitive organelle, organizing reactions that require special conditions for optimization, such as the sequestration of substrates, cofactors, or toxic intermediates.

Carboxysomes can be grouped into two classes, α and β , based on sequence analysis and gene organization. In β -carboxysomes, the shell proteins are the *ccmK* gene products (named for carbon-concentrating mechanism) (11, 12, 16). Several species of cyanobacteria including *Synechocystis* sp. PCC 6803 contain four similar genes coding for the CcmK proteins, CcmK1 to 4 (Fig. 2). We have determined the crystal structures of CcmK2 and CcmK4 (in two crystal forms), to reveal the structural fold of the widely distributed BMC domain. The overall three-dimensional folds of CcmK4 and CcmK2 are practically identical (Fig. 3, A and B). Searches of the databases of known three-dimensional structures (17, 18) reveal that the N-terminal 80 amino acids of the BMC domain adopt an α/β fold, observed in ferredoxin and numerous apparently unrelated proteins. No direct evolutionary link could be established either by sequence or structure comparisons, between the BMC domain and known viral capsid proteins.

In all three structures visualized, the protein subunits are arranged in hexameric units about a central six-fold axis of symmetry (Fig. 3B). The structure of the BMC domain monomer is notably wedge-shaped, so that six subunits fit together to leave only a small central hole. Monomers of a more spherical shape would assemble to leave a circular hole about the size

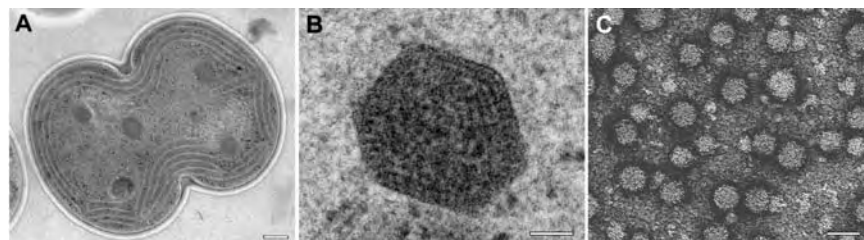
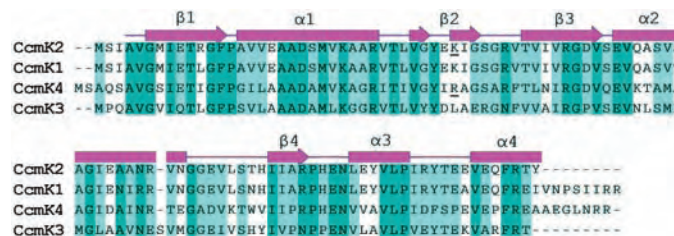


Fig. 1. (A and B) Transmission electron micrographs (EMs) of cyanobacterium *Synechocystis* sp. PCC 6803 showing (A) a single cell dividing (scale bar, 200 nm). The polyhedral shape and crystalline structure distinguishes the carboxysomes from other cytoplasmic inclusions. (B) A single carboxysome (scale bar, 50 nm). (C) Transmission EMs of purified recombinant BMC domain protein CcmK2, after precipitation and sonication (scale bar, 50 nm). EMs in (A) and (B) provided by A. M. L. van de Meene, W. F. J. Vermaas, and R. W. Roberson, School of Life Sciences, Arizona State University.

Fig. 2. A sequence alignment of the four BMC domain proteins in *Syn* 6803, CcmK1 to 4. Regions of high conservation are shaded, and the protein's secondary structure elements as determined here are shown (rectangles, α helices; arrows, β strands). Charged amino acids found in the hexamer pore (Fig. 3, C and D) are underlined.



¹Molecular Biology Institute, University of California, Los Angeles (UCLA), Box 951570, ²Life Sciences Core, UCLA, Box 160606, ³Department of Chemistry and Biochemistry, UCLA, Box 951569, Los Angeles, CA 90095, USA, ⁴UCLA–U.S. Department of Energy (DOE) Institute for Genomics and Proteomics, Box 951570, Los Angeles, CA 90095–1570, USA.

*To whom correspondence should be addressed. E-mail: yeates@mbi.ucla.edu

of one subunit. The pore at the center of the CcmK2 hexamer (taking into account atomic van der Waals radii) is ~ 7 Å in diameter (Fig. 3C). The CcmK4 pore is narrower, ~ 4 Å in diameter (Fig. 3D). The narrowness of the hole in the BMC domain hexamer apparently reflects a functional constraint. A further unexpected finding is a large net positive electrostatic potential within the central pore. Each of the six subunits contributes a conserved positively charged amino acid residue (Arg³⁸ in CcmK4; Lys³⁶ in CcmK2) to the pore (Fig. 3, C and D). The energetic cost of bringing together such a large number of like charges is considerable (SOM text). We conclude, from the cost of creating such a configuration and from its conservation in numerous carboxysome shell proteins from different organisms, that the charged pore serves a functional role in the carboxysome, most likely by regulating metabolite flux.

In the crystal structure of CcmK2 and in the second crystal form of CcmK4, the next level of subunit organization is evident (Fig. 4). Owing to their nearly flat edges, the hexameric units fit together readily. For CcmK2, the hexameric units form a nearly solid molecular sheet (Fig. 4A). In the second crystal form of CcmK4, we observe linear strips of hexamers (Fig. 4B). Two adjacent CcmK4 hexamers can be superimposed on two adjacent CcmK2 hexamers with only a 1.4 Å root mean square difference over 551 Ca atoms. The arrangement of hexamers in the CcmK2 sheet leaves only small gaps at the two-fold and three-fold axes of symmetry in the layer, where two and three hexameric units come together. The close packing of the hexamers in sheets is in marked contrast to other nonviral hexameric protein assemblies that have been observed (19). On the basis of

the recurrence of the same interface in different crystal forms, involving distinct subunit types and of the unusually tight packing achieved, we argue that the observed packing arrangement is biologically relevant. The sheet structure suggests that the carboxysome shell is roughly 2 to 3 nm thick. According to the spacing of subunits in the hexagonal sheet, a microcompartment with a diameter of 150 nm (Fig. 1B) would contain some 10,000 BMC domain proteins in its outer shell.

In addition to the hexameric pore, the gaps between hexamers may serve as conduits for metabolites. The gap at the three-fold axis is ~ 6 Å in diameter (if one takes into account atomic radii), whereas the gap at the two-fold axis is elongated, but only ~ 4 Å wide. Charged amino acid residues, conserved across CcmK proteins, are also notable at these gaps (SOM text). Especially in view of possible minor side chain rearrangements, the pores and gaps appear to be large enough to allow transit of the negatively charged metabolites (namely, bicarbonate, ribulose bis-phosphate, phosphoglycerate, and possibly hydroxyl ions) that have to cross the carboxysome shell. In contrast, uncharged molecules such as CO₂ and O₂ would not be attracted to the charged pores and gaps. The possibility that pores within and gaps between hexamers might serve as portals for small molecules is reminiscent of certain viruses in which it has been determined that the interstitial spaces between subunits mediate the flow of metabolites into and out of the capsids (20, 21).

The structural findings suggest that the carboxysome shell is involved in controlling metabolite flow, rather than serving merely as a storage device for excess RuBisCO or simply to localize the requisite enzymes for carbon fixation. Such selective permeability is fundamental to subcellular organelles. Although bacterial microcompartments do not meet the traditional criteria (for eukaryotic organelles) of being membrane-bound, the view emerging here is that they meet the requirements of organelle function in other ways, for example by controlling flow through a nearly solid protein shell rather than through proteins embedded in a continuous membrane.

When they were first observed in bacterial cells, the resemblance of carboxysomes to viral particles was so striking that it prompted efforts to induce viral-mediated lysis of the cells (4). Not surprisingly, our structural studies reveal a striking number of parallels between the basic features of bacterial microcompartments and viruses (22–27). Both are highly symmetrical structures sharing underlying design features. In some viruses, the subunit organization is best described by combinations of hexameric and pentameric capsomeres that self-assemble. According to the rules of solid geometry, forming a closed shell generally requires 12 pentameric capsomeres to

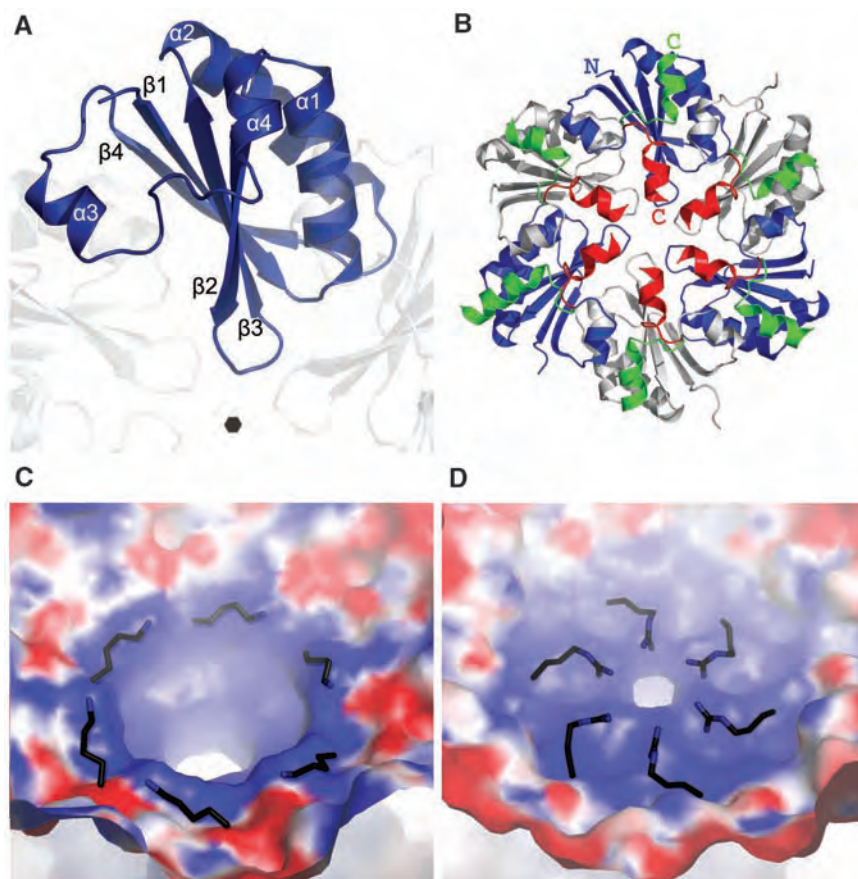
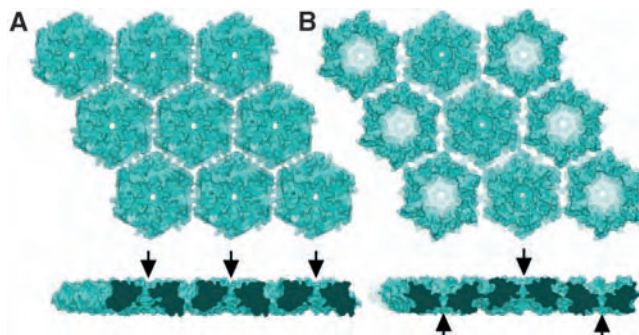


Fig. 3. (A) The three-dimensional crystal structure of BMC domain protein, CcmK4, determined at a resolution of 1.8 Å. (B) The concave surface of the CcmK4 hexamer. CcmK4 monomers are colored alternately blue and gray, with all six C termini in green. Also shown is the CcmK2 C terminus (red) in its corresponding position after a superposition of CcmK2 and CcmK4 hexamers. The superposition (not shown) indicates that the backbones of the two hexamers are nearly identical up to residues Pro⁹⁷ (CcmK4)/Glu⁹⁵ (CcmK2). The structural differences at the C termini (labeled) suggest one reason why CcmK4 packing is limited to chains of hexamers. In CcmK4, the C termini extend outward, toward the corners of the hexamer. These are the positions where three hexamers meet in the CcmK2 sheet but not in the CcmK4 structures. (C and D) A close-up view of the pores (concave side) formed at the six-fold axis of symmetry in the CcmK2 (C) and CcmK4 (D) hexamers. The surfaces are colored according to electrostatic potential, with blue positive and red negative. Positively charged, conserved amino acid side chains lining the pore are highlighted (Lys³⁶ in CcmK2, Arg³⁸ in CcmK4). This figure and Fig. 4 were illustrated with PyMOL (32).

Fig. 4. Crystal packing of BMC domain proteins in molecular layers. (A) CcmK2 hexamers packed in uniform orientations (convex face shown). (B) CcmK4 hexamers (crystal form 2) packed in strips of alternating orientation. In the side view of each sheet, arrows mark the positions of the pores.



be present among a greater number of hexameric capsomeres (SOM text). These ideas are consistent with our studies of the BMC domain proteins. The first three crystal structures have revealed hexameric assemblies; sedimentation and native gel electrophoresis studies indicate that some of the subunits form pentamers in addition to hexamers (fig. S2, SOM text).

Likewise, similar but distinct protein subunits existing in quasi-equivalent environments or forms are required to construct virus capsids (22–27). Among the CcmK proteins, CcmK2 alone appears capable of forming closed shells under certain conditions, but these structures are much smaller than native carboxysomes and lack their polyhedral regularity (Fig. 1C). Apparently, as with many viruses, multiple distinct carboxysome subunit types appear to be required to achieve the correct architecture. The generally high conservation of amino acid residues among the CcmK paralogs at the interhexamer interfaces is consistent with the idea that hexameric (or possibly pentameric) units of different types could assemble together. The subtle differences in primary and tertiary structure could then govern the quasi-equivalent interactions necessary to create the native shell. The crystal structures of CcmK2 and CcmK4 give preliminary clues as to the origins of distinct assembly behavior, such as their disparate tendency to form sheets (Fig. 4). In CcmK4, clashes between C termini from adjacent hexamers could preclude the formation of flat sheets of the type seen with the CcmK2 hexamer. Interactions between the C-terminal tails of the BMC domain proteins could influence microcompartment domain assembly in the same way that the flexible termini of certain viral capsid proteins often participate as switches for distinct types of interactions in the mature viral capsid (25, 26, 28, 29). The shape and assembly of the carboxysome could also be affected by other proteins that may be present in the shell (30) whose structures are not yet known.

As was the case for other large molecular machines such as viruses and ribosomes (31), fully elucidating the structure and function of bacterial microcompartments will require combining electron microscopy, biophysical experiments, and crystallographic studies on

the multiple components in order to attain an understanding of the whole.

References and Notes

1. T. E. Jensen, C. C. Bowen, *Proc. Iowa Acad. Sci.* **68**, 89 (1961).
2. E. Gantt, S. F. Conti, *J. Bacteriol.* **97**, 1486 (1969).
3. J. M. Shively, G. L. Decker, J. W. Greenawalt, *J. Bacteriol.* **96**, 2138 (1970).
4. J. M. Shively, R. S. English, *Can. J. Bot.* **69**, 957 (1991).
5. J. M. Shively, F. Ball, D. H. Brown, R. E. Saunders, *Science* **182**, 584 (1973).
6. G. D. Price, J. R. Coleman, M. R. Badger, *Plant Physiol.* **100**, 784 (1992).
7. A. K.-C. So, G. S. Espie, *Plant Mol. Biol.* **37**, 205 (1998).
8. A. K.-C. So *et al.*, *J. Bacteriol.* **186**, 623 (2004).
9. G. A. Codd, W. J. N. Marsden, *Biol. Rev. Camb. Philos. Soc.* **59**, 389 (1984).
10. G. D. Price, M. R. Badger, *Can. J. Bot.* **69**, 963 (1990).
11. G. C. Cannon *et al.*, *Appl. Environ. Microbiol.* **67**, 5351 (2001).
12. G. C. Cannon, S. Heinhorst, C. E. Bradburne, J. M. Shively, *Funct. Plant Biol.* **29**, 175 (2002).
13. J. M. Shively *et al.*, *Can. J. Bot.* **76**, 906 (1998).
14. E. Kofoed, C. Rappleye, I. Stojiljkovic, J. Roth, *J. Bacteriol.* **181**, 5317 (1999).
15. G. D. Havemann, T. A. Bobik, *J. Bacteriol.* **185**, 5086 (2003).

16. M. R. Badger, D. Hanson, G. D. Price, *Funct. Plant Biol.* **29**, 161 (2002).
17. L. Holm, C. Sander, *J. Mol. Biol.* **233**, 123 (1993).
18. J.-F. Gibrat, T. Madej, S. H. Bryant, *Curr. Opin. Struct. Biol.* **6**, 377 (1996).
19. P. Yuan *et al.*, *Nat. Struct. Biol.* **10**, 264 (2003).
20. J. M. Diprose *et al.*, *EMBO J.* **20**, 7229 (2001).
21. H. Naitow, J. Tang, M. Canady, R. B. Wickner, J. E. Johnson, *Nat. Struct. Biol.* **9**, 725 (2002).
22. D. L. D. Caspar, A. Klug, *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1 (1962).
23. J. M. Hogle, M. Chow, D. J. Filman, *Sci. Am.* **256**, 42 (March 1987).
24. M. G. Rossmann, J. E. Johnson, *Annu. Rev. Biochem.* **58**, 533 (1989).
25. S. C. Harrison, *Curr. Opin. Struct. Biol.* **5**, 157 (1995).
26. J. E. Johnson, J. A. Speir, *J. Mol. Biol.* **269**, 665 (1997).
27. S. C. Harrison, *Curr. Opin. Struct. Biol.* **11**, 195 (2001).
28. T. Stehle, S. J. Gamblin, Y. Yan, S. C. Harrison, *Structure* **4**, 165 (1996).
29. C. San Martin *et al.*, *Nat. Struct. Biol.* **9**, 756 (2002).
30. A. K. So *et al.*, *J. Bacteriol.* **186**, 623 (2004).
31. N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, *Science* **289**, 905 (2000).
32. W. L. DeLano, www.pymol.org (2002).
33. We thank G. Cannon and M. Yeager for helpful discussions, D. Krogmann for the gift of *Synechocystis* sp. PCC 6803 genomic DNA, and J. Laidman and H. Adesiteyo for technical assistance. We also thank R. W. Roberson, W. F. J. Vermass, and A. van de Meene for providing transmission electron micrographs of *Syn* 6803 cells. C.A.K. and T.O.Y. acknowledge the U.S. Department of Agriculture, the National Institutes of Health, and the Office of Science Biological and Environmental Research Program, the U.S. Department of Energy. The coordinates and structure factors have been deposited in the Protein Data Bank with accession numbers 2A10 (CcmK4, Form 1); 2A18 (CcmK4, Form 2); and 2A1B (CcmK2).

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5736/936/DC1

Materials and Methods

SOM Text

Figs. S1 to S3

References and Notes

11 April 2005; accepted 22 June 2005
10.1126/science.1113397

Rewiring of the Yeast Transcriptional Network Through the Evolution of Motif Usage

Jan Ihmels,¹ Sven Bergmann,¹ Maryam Gerami-Nejad,² Itai Yanai,¹ Mark McClellan,² Judith Berman,² Naama Barkai^{1*}

Recent experiments revealed large-scale differences in the transcription programs of related species, yet little is known about the genetic basis underlying the evolution of gene expression and its contribution to phenotypic diversity. Here we describe a large-scale modulation of the yeast transcription program that is connected to the emergence of the capacity for rapid anaerobic growth. Genes coding for mitochondrial and cytoplasmic ribosomal proteins display a strongly correlated expression pattern in *Candida albicans*, but this correlation is lost in the fermentative yeast *Saccharomyces cerevisiae*. We provide evidence that this change in gene expression is connected to the loss of a specific cis-regulatory element from dozens of genes following the apparent whole-genome duplication event. Our results shed new light on the genetic mechanisms underlying the large-scale evolution of transcriptional networks.

Evolution of gene expression plays a prominent role in generating phenotypic diversity (1–3), but little is known about the genetic

basis underlying broad modulations of the genome-wide transcription program. Here we describe a rewiring of the yeast transcrip-

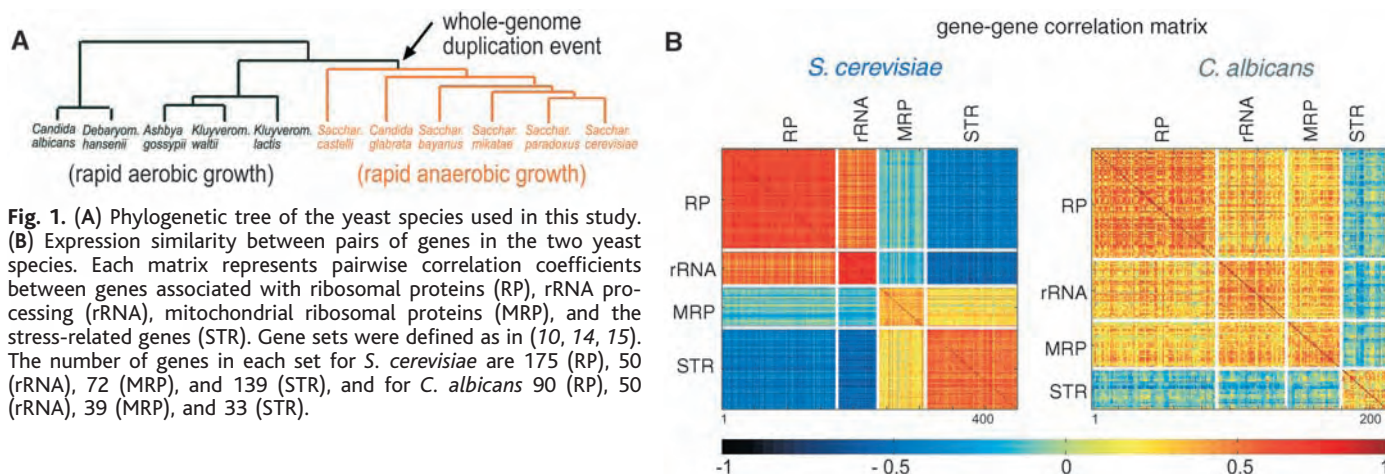


Fig. 1. (A) Phylogenetic tree of the yeast species used in this study. **(B)** Expression similarity between pairs of genes in the two yeast species. Each matrix represents pairwise correlation coefficients between genes associated with ribosomal proteins (RP), rRNA processing (rRNA), mitochondrial ribosomal proteins (MRP), and the stress-related genes (STR). Gene sets were defined as in (10, 14, 15). The number of genes in each set for *S. cerevisiae* are 175 (RP), 50 (rRNA), 72 (MRP), and 139 (STR), and for *C. albicans* 90 (RP), 50 (rRNA), 39 (MRP), and 33 (STR).

tional network that is linked to the massive loss of a conserved cis-regulatory element.

With the recent sequencing of many yeast genomes, the hemiascomycete lineage emerged as a central paradigm for studying the genetic basis of phenotypic diversity (4). Despite a strong conservation of gene content, yeast species exhibit major phenotypic differences. Most striking, perhaps, is their differential requirement for oxygen. Whereas growth of most yeast species requires oxygen, *Saccharomyces cerevisiae* grows rapidly in its absence and prefers to ferment glucose anaerobically even when oxygen is present. The emergence of this capacity is linked to the apparent whole-genome-duplication event (5–7), which allowed the specialization of enzymes to aerobic versus hypoxic conditions (4, 8) (Fig. 1A).

The preference for fermentation is reflected in the organization of the *S. cerevisiae* transcriptional network. For example, genes involved in oxidative phosphorylation are repressed in the presence of glucose (9). To examine the transcription program of *S. cerevisiae* in more detail, we compiled a database of over 1000 published genome-wide expression profiles (10). In accordance with its anaerobic growth phenotype, we found that the expression of *S. cerevisiae* genes encoding mitochondrial ribosomal proteins (MRP), is not correlated (and perhaps is even weakly anticorrelated) with that of genes coding for cytoplasmic ribosomal proteins (RP) and with ribosomal RNA (rRNA) processing genes (10) (Fig. 1B). Instead, expression of the MRP genes exhibits a distinct correlation with that of genes induced in response to environmental stress conditions. At least in part, this correlation reflects the induction of the stress-related genes during the

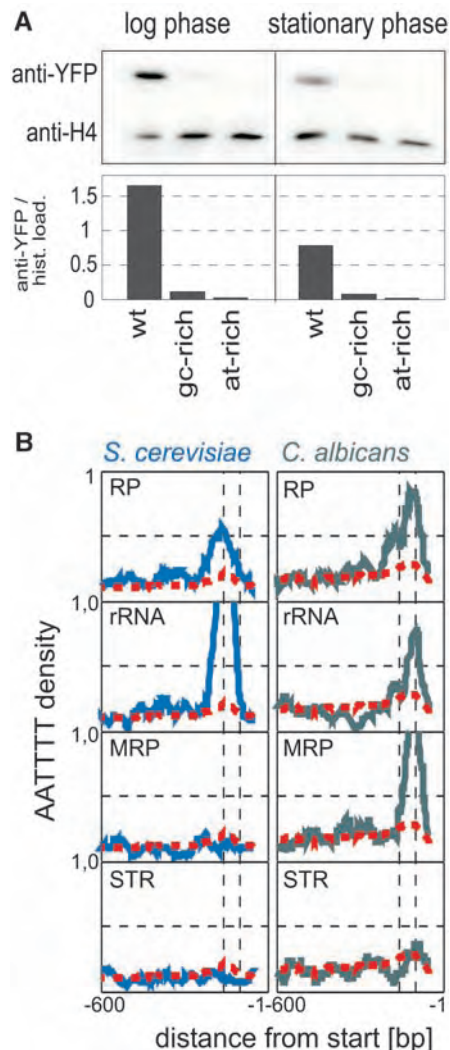


Fig. 2. (A) The AATTTT sequence is required to induce high-level transcription of a reporter protein fused to the *C. albicans* MRP7 promoter in exponentially growing cells. **(B)** Frequency of occurrence of RGE in a 50-bp window of the gene groups indicated. Background genomic frequency is indicated in red (15). Multiple occurrences of the RGE were included. The same quantitative result is obtained when only a single occurrence per gene is considered (fig. S2).

¹Department of Molecular Genetics and Department of Physics of Complex systems, Weizmann Institute of Science, Rehovot, Israel. ²Department of Genetics, Cell Biology and Development and Department of Microbiology, University of Minnesota, MN 55455, USA.

*To whom correspondence should be addressed. E-mail: naama.barkai@weizmann.ac.il

relatively slow growth in nonfermentable carbon sources, such as glycerol or ethanol, which requires mitochondrial function.

To examine the manifestation of this relationship between the expression of MRP genes and genes required for rapid growth in aerobic yeast species, we assembled a data set of 198 genome-wide expression profiles of *Candida al-*

bicans, a human pathogen that primarily grows aerobically. In accordance with the requirement for respiration to support rapid growth, we found that in *C. albicans* the expression of MRP genes is strongly correlated with the expression of genes involved in ribosome assembly, including RP genes and rRNA processing genes (Fig. 1B). Thus, the organization of the *S. cerevisiae*

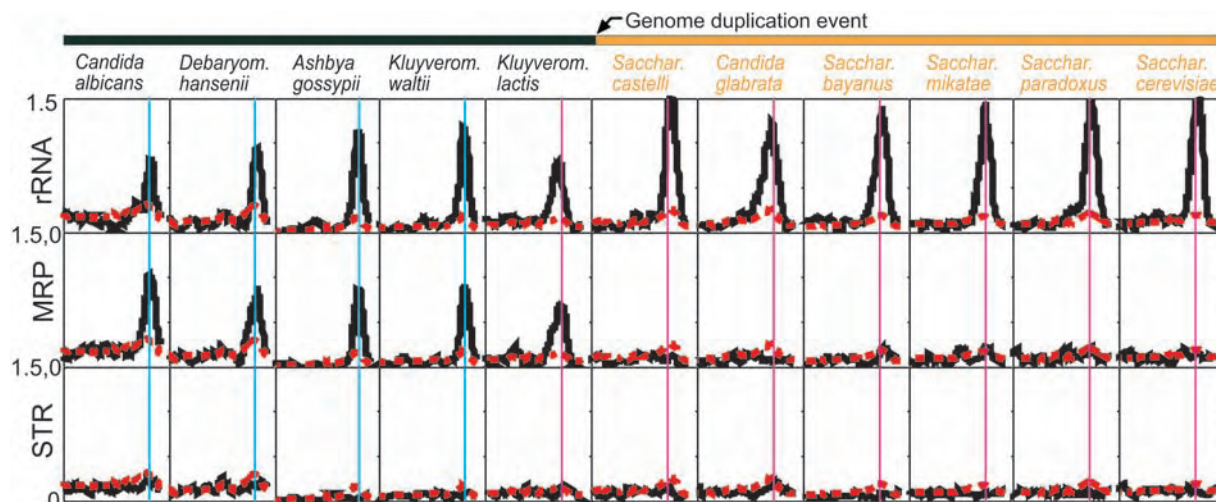


Fig. 3. Frequency of occurrence of RGE in intermediate yeast species (15). The gene sets for *S. cerevisiae* were defined as in Fig. 1B. Gene sets for the remaining species consist of orthologs of the *S. cerevisiae* genes.

and *C. albicans* transcriptional networks reflects their respective physiologies.

Gene expression can evolve through mutations in trans-acting regulatory proteins or in cis-acting DNA elements (compare fig. S4). For example, the regulation of a common transcription factor could diverge between the species. Alternatively, modifications of the individual cis-regulatory elements of each downstream gene could change their regulation, potentially coopting or losing regulation by a transcription factor. To distinguish between these scenarios, we searched the regions 600 base pairs (bp) upstream of open reading frames of both organisms for overrepresented sequence elements. Specific regulatory sequences that appear to be conserved between *S. cerevisiae* and *C. albicans* were identified (11, 12). Most notably, the PAC (GATGAG) sequence (13) was associated with rRNA processing genes in both species. In agreement with previous reports (10), no overrepresented sequence was associated with the MRP genes in *S. cerevisiae*.

In contrast, in *C. albicans*, the MRP genes were clearly associated with an overrepresented upstream sequence motif (AATTTT). This sequence was previously implicated in the regulation of rRNA processing genes in *S. cerevisiae* (13), although its functional role was not demonstrated experimentally. To examine whether this motif contributes to the regulation of MRP genes in *C. albicans*, we fused the promoter of the MRP7 gene to a yellow fluorescent protein (YFP) reporter. As expected, strong induction of the YFP reporter was observed during exponential growth, whereas expression in stationary phase was much weaker (Fig. 2A). Mutating the AATTTT sequence in the MRP7 promoter to either a GC-rich or to an AT-rich sequence drastically reduced expression of the YFP (Fig. 2A), indicating that it is the AATTTT sequence itself, and not its AT-rich nature, that is important for its function. Similar results were obtained with two addi-

tional mitochondrial ribosomal protein gene promoters (fig. S5). We conclude that the AATTTT sequence functions as a cis-regulatory element of MRP genes in *C. albicans*.

Examining the appearance of this motif in more detail, we found that its position relative to that of the ORF start codon is highly confined in both *S. cerevisiae* and *C. albicans*, although the precise position is somewhat different (160 versus 110 bp in *S. cerevisiae* and *C. albicans*, respectively). In both organisms, this sequence is significantly overrepresented in genes involved in rRNA processing and ribosomal proteins, but not in genes associated with the environmental stress response. In addition, only in *C. albicans* is it overrepresented also in MRP genes (Fig. 2B). Because this conserved element is associated with genes required for rapid growth in both organisms, we denote it as a rapid growth element (RGE).

This RGE could have been coopted to *C. albicans* MRP genes, or alternatively, could have been lost in *S. cerevisiae*. To distinguish between these alternatives, we examined the appearance of the RGE in nine sequenced yeast species that are considered to be intermediate in the evolution between *S. cerevisiae* and *C. albicans* (Fig. 3). In all species examined, the RGE is significantly overrepresented in genes involved in rRNA processing. Notably, overrepresentation of the RGE in MRP promoters was found in all genomes that diverged from the *S. cerevisiae* lineage before the whole-genome duplication event (Fig. 3). This indicates a widespread loss of the RGE in MRP genes following the genome duplication.

Taken together, it appears that the emergence of anaerobic growth capacity in yeast is associated with a global rewiring of the yeast transcriptional network. This rewiring involved changes in the promoter regions of dozens of genes, manifested by the loss of a specific regulatory motif from MRP gene promoters. It would be interesting to examine which tran-

scription factor binds to the RGE site and whether its coding sequence was also modified in association with the emergence of anaerobic growth. Intriguingly, the most parsimonious scenario for the described promoter adaptation appears to be intimately linked with the whole-genome duplication event. Our results suggest that gene duplication can facilitate the evolution of new function not only by specialization of coding sequences but also by facilitating the evolution of gene expression.

References and Notes

1. M. C. King, A. C. Wilson, *Science* **188**, 107 (1975).
2. S. B. Carroll, *Cell* **101**, 577 (2000).
3. E. H. Davidson, *Genomic Regulatory Systems: Development and Evolution* (Academic Press, San Diego, CA, 2001).
4. J. Piskur, R. B. Langkjaer, *Mol. Microbiol.* **53**, 381 (2004).
5. K. H. Wolfe, D. C. Shields, *Nature* **387**, 708 (1997).
6. M. Kellis, B. W. Birren, E. S. Lander, *Nature* **428**, 617 (2004).
7. F. S. Dietrich et al., *Science* **304**, 304 (2004).
8. K. Wolfe, *Curr. Biol.* **14**, R392 (2004).
9. H. J. Schuller, *Curr. Genet.* **43**, 139 (2003).
10. J. Ihmels et al., *Nat. Genet.* **31**, 370 (2002).
11. A. P. Gasch et al., *PLoS Biol.* **2**, 398 (2004).
12. A. Tanay, A. Regev, R. Shamir, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7203 (2005).
13. S. Tavazoie, G. M. Church, *Nat. Biotechnol.* **16**, 566 (1998).
14. S. Bergmann, J. Ihmels, N. Barkai, *PLoS Biol.* **2**, E9 (2004).
15. Methods are available as supporting material on Science Online. See also www.weizmann.ac.il/home/barkai/Rewiring.
16. We thank the following for providing transcription profiling data sets: R. Bennett, A. Tsong, A. Johnson, M. Lorenz, C. D'enfert, G. Fink, M. Whiteway, A. Nantel, P. D. Rogers, and especially D. Davis, D. Marcus, B. Hube, and D. Sanglard, for providing transcription profile data sets before publication. We also thank A. Nantel, C. D'enfert, and A. Tsong for providing gene name information that allowed assignment of genes on different arrays to their orf19 gene identities. This work was supported by NIH grants A150562 (N.B.) and R01 DE/AL14666 (J.B.) and a grant from the Kahn Fund for Systems Biology at the Weizmann Institute of Science (N.B.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5736/938/DC1
Methods
Figs. S1 to S5
Tables S1 to S4

20 April 2005; accepted 23 June 2005
10.1126/science.1113833

Export-Mediated Assembly of Mycobacterial Glycoproteins Parallels Eukaryotic Pathways

Brian C. VanderVen, Jeffery D. Harder, Dean C. Crick, John T. Belisle*

Protein *O*-mannosylation is an essential and evolutionarily conserved post-translational modification among eukaryotes. This form of protein modification is also described in *Mycobacterium tuberculosis*; however, the mechanism of mannoprotein assembly remains unclear. Evaluation of differentially translocated chimeric proteins and mass spectrometry to monitor glycosylation demonstrated that specific translocation processes were required for protein *O*-mannosylation in *M. tuberculosis*. Additionally, Rv1002c, a *M. tuberculosis* membrane protein homolog of eukaryotic protein mannosyltransferases, was shown to catalyze the initial step of protein mannosylation. Thus, the process of protein mannosylation is conserved between *M. tuberculosis* and eukaryotic organisms.

Like eukaryotes, prokaryotes modify proteins with a variety of glycosyl residues that influence a number of biological events (1, 2). In *Mycobacterium tuberculosis* (*Mtb*), two fully characterized glycoproteins (Apa/Rv1860 and MPB83/Mb2898) possess threonine residues with linear $\alpha(1\rightarrow2)$ and $\alpha(1\rightarrow3)$ oligomannosides, a glycosylation pattern reminiscent of eukaryotic short-chain mannoproteins (3–5). Other proteins of *Mtb* are known to be glycosylated, presumably with mannose residues as defined by ConcanavalinA (ConA) reactivity (6, 7). As with other bacterial pathogens, the glycosylation of mycobacterial proteins influences host interactions; for example, antigen-specific T cell recognition of Apa requires mannosylation (8, 9), and LpqH, an *O*-mannosylated lipoprotein, is a Toll-like receptor 2 (TLR-2) agonist (10).

The spatial assembly of a few bacterial glycoproteins has been addressed (11–13), and

although most bacterial glycoproteins are extracytoplasmic (1), little evidence exists to link glycosylation with any of the bacterial protein secretion pathways. In *Mtb*, the biosynthetic pathway of protein glycosylation remains largely unknown. However, similarities exist between the *O*-mannosylation of *Mtb* proteins and protein *O*-mannosylation in *Saccharomyces cerevisiae*. Specifically, mycobacterial and yeast mannoproteins are exported in a Sec-dependent fashion (3, 4, 6, 7), and protein *O*-mannosyltransferase (PMT) activity of both organisms is membrane associated, requiring a lipid carrier to donate mannose (5, 14). These similarities between otherwise disparate organisms suggest that the enzymatic machinery for protein *O*-mannosylation might also be conserved. The PMTs of *S. cerevisiae* are integral membrane proteins sharing amino acid sequence identities of 57.5% (5). Bioinformatic approaches identified a single *Mtb* PMT homolog, designated Rv1002c. This putative 55.5-kD protein displayed 22 to 24% identity with the PMTs of *S. cerevisiae* and presented a similar hydropathy profile (fig. S1). More prognostic of potential PMT activity were two conserved Arg residues in transmembrane domains and an Asp-

Glu motif in the first predicted extracytoplasmic domain of Rv1002c (Fig. 1A). The Arg and Glu residues are essential and form the putative active site in yeast PMTs (15). These invariant residues are also common to the PMTs of higher eukaryotes such as humans (16).

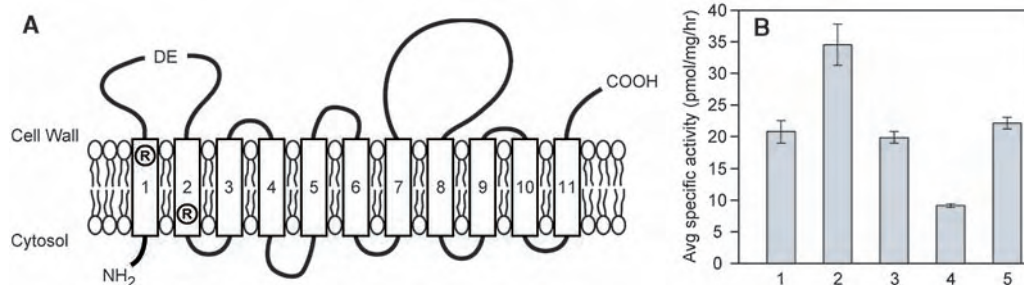
To demonstrate PMT activity, wild-type Rv1002c and three mutated forms of this gene conferring A₅₅A₅₆, A₅₅E₅₆, or D₅₅A₅₆ substitutions in the invariant D₅₅E₅₆ motif were overexpressed in *Mycobacterium smegmatis*, and membrane preparations were assayed for in vitro PMT activity. Overexpressed wild-type *Mtb* Rv1002c significantly increased (66.8%) PMT activity of the *M. smegmatis* membranes above that observed in the vector control (Fig. 1B). Only wild-type recombinant Rv1002c conferred increased PMT activity and, similar to the PMTs of *S. cerevisiae*, the invariant D₅₅E₅₆ motif of Rv1002c was critical to proper function (Fig. 1B). In particular, membranes containing recombinant Rv1002c with a D₅₅E₅₆ to A₅₅E₅₆ substitution resulted in decreased (56.0%) PMT activity compared with the *M. smegmatis* vector control. Thus, altered forms of Rv1002c apparently competed for substrate binding without catalysis of glycosylation.

The membrane location of the *Mtb* PMT (Rv1002c) and Sec-translocation of *Mtb* mannoproteins suggested a link between protein export and glycosylation. This hypothesis was partially supported by the reactivity of ConA to a subset of membrane and culture filtrate (CF) proteins but not cytosolic proteins (fig. S2). To evaluate the association between protein *O*-mannosylation and Sec-translocation, a recombinant form of FbpC/Rv0129c, a well-characterized Sec-translocated and nonglycosylated protein of *Mtb* (17), was fused with and without its signal peptide (SP) to a mannosylation cassette (MC) (DPEAPPVPTTAASP) of Apa (3) and a hexa-His tag (Fig. 2A). These two chimeric proteins differentially localized on the basis of the presence or absence of a Sec-dependent translocation SP. SP(+)-FbpC-MC localized to the membrane, cell wall (CW), and CF, but not the cytosol, of *Mtb* and *M. smegmatis* (Fig. 2B). N-terminal sequencing of the CW and CF

Mycobacteria Research Laboratories, Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, CO 80523–1682, USA.

*To whom correspondence should be addressed. E-mail: jbelisle@colostate.edu

Fig. 1. Membrane topology of Rv1002c and in vitro protein mannosylation activity. (A) Bioinformatic analyses predicted 11 transmembrane domains in Rv1002c and the presence of the conserved PMT active site residues R₃₉, D₅₅, E₅₆, and R₁₀₆. (B) In vitro PMT activity of cytoplasmic membrane preparations from log-phase recombinant *M. smegmatis* clones as determined through a composite assay that assesses the transfer of mannose from guanosine diphosphate mannose through mannosyl-phosphoryl-decaprenol (14) to a synthetic peptide (AAAPPATPVAPPPHHHHHH) encompassing a single glycosylation domain of Apa. 1, vector control; 2, wild-type Rv1002c; 3, Rv1002c DE→AA mutant; 4, Rv1002c DE→AE mutant; 5,



Rv1002c DE→DA mutant. Error bars indicate SD of three replicate experiments. Reverse transcription polymerase chain reaction analyses confirmed the expression of the recombinant forms of Rv1002c in *M. smegmatis*.

forms of SP(+)-FbpC-MC revealed the sequence (FSRPGLPVEY) of the mature amino terminus of FbpC (17), confirming that SP(+)-FbpC-MC was processed through the Sec-translocase. In comparison, SP(-)-FbpC-MC was observed in the cytosol of *M. smegmatis* as expected but was also found in the membrane and CW of this bacterium and was predominantly in the CW of *Mtb*, with a truncated form in the membrane (Fig. 2B). The

occurrence of SP(-)-FbpC-MC in membrane and CW fractions was probably due to mixing of protein pools during subcellular fractionation and the fatty acid binding domain of FbpC (18). Indeed, the absence of SP(-)-FbpC-MC from the CF demonstrated that this construct was not actively secreted. Instability of SP(-)-FbpC-MC in *Mtb* probably accounted for its absence from the cytosol and the apparent molecular weight shift in the membrane fraction.

The glycosylation status of *Mtb*- and *M. smegmatis*-derived SP(+)-FbpC-MC and SP(-)-FbpC-MC purified from the CF and CW was determined after affinity chromatography purification. SP(+)-FbpC-MC was ConA reactive; however, SP(-)-FbpC-MC did not bind ConA (fig. S3). Liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MS/MS) of the trypsin-digested *Mtb* SP(+)-FbpC-MC yielded three major $[M+H]^{3+}$ molecular ions that differed by 54 m/z (mass/charge ratio) (Fig. 3A) and corresponded to the MC with zero, one, and two hexose residues. MS/MS fragmentation of the largest molecular ion (1327.0 m/z) generated the $[M+H]^{3+}$ daughter ions of 1273.7 and 1219.6 m/z arising from neutral losses of 54 m/z (Fig. 3B), a diagnostic property for the dissociation of hexose residues from a triply charged precursor peptide (3). The relative abundance of parent ions demonstrated in a semiquantitative manner that the majority of SP(+)-FbpC-MC was glycosylated, with the most dominant form having a single hexose (Fig. 3A). Identical experiments with the *M. smegmatis* forms of this protein demonstrated similar patterns of glycosylation. In contrast, LC-ESI-MS/MS analysis of trypsin-digested SP(-)-FbpC-MC recovered from the *Mtb* CW and from the *M. smegmatis* cytosol and CW revealed a single $[M+H]^{3+}$ molecular ion (1218.9 m/z) that corresponded to the non-glycosylated MC (Fig. 3C). Thus, only fusion constructs with the Sec-dependent SP were presented in a manner that allowed glycosylation.

Mtb produces several proteins that are secreted in a Sec-independent manner. Two such proteins, ESAT-6/Rv3875 and SodA/Rv3846, appear in the CF during early log-phase growth

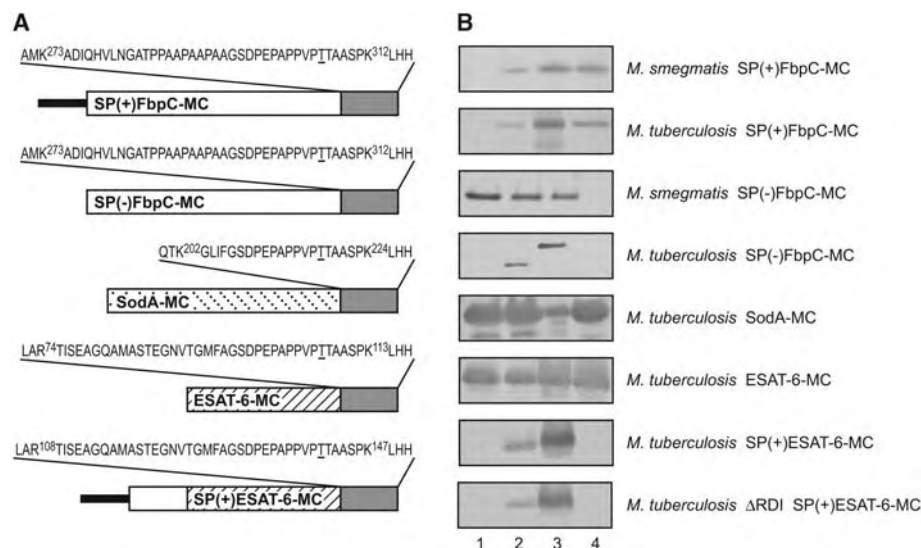


Fig. 2. Protein chimera structures and their localization in *Mycobacterium* spp. (A) Schematic representation of protein chimeras generated to evaluate the relation between O-mannosylation and translocation. Amino acid sequences shown indicate the tryptic-generated fragment containing the Apa MC, with the known glycosylation site (3) underlined (26). The heavy dark line depicts the SP of FbpC that targets proteins for Sec-translocation, the open box depicts the mature FbpC, the gray box represents the Apa MC and hexa-His tag, the dotted box depicts SodA, and the slashed box depicts ESAT-6. (B) Subcellular localization of each chimera produced in *Mtb* and *M. smegmatis* was determined by Western blotting using a monoclonal antibody to His₆ as the probe. Lane 1, cytosol; lane 2, cell membrane; lane 3, CW; lane 4, CF.

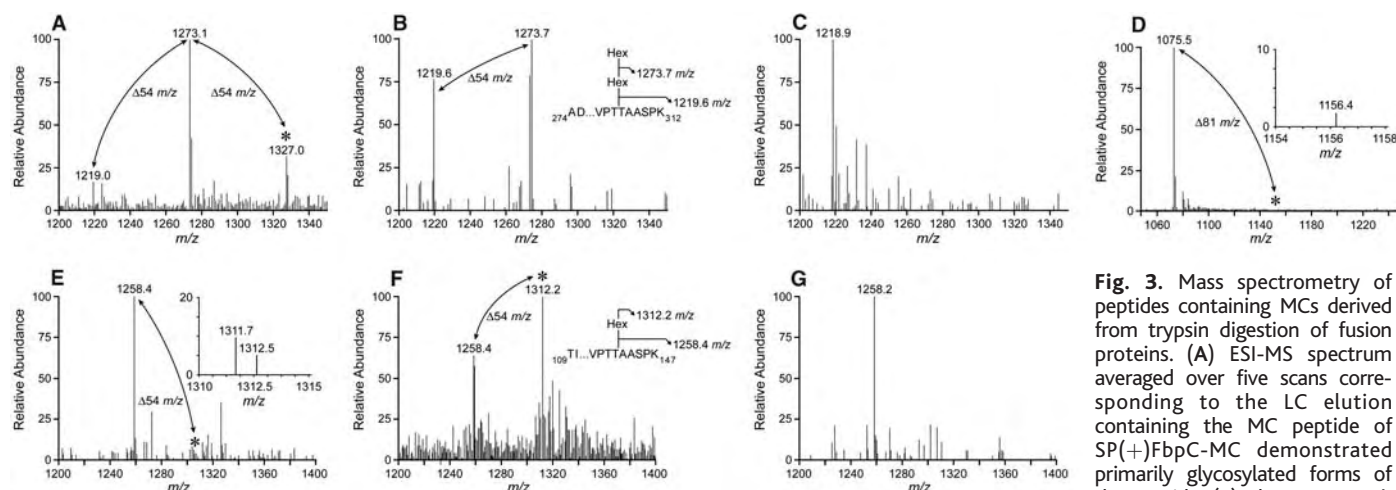


Fig. 3. Mass spectrometry of peptides containing MCs derived from trypsin digestion of fusion proteins. (A) ESI-MS spectrum averaged over five scans corresponding to the LC elution containing the MC peptide of SP(+)-FbpC-MC demonstrated primarily glycosylated forms of the peptide. (B) The 1327.0 m/z

$[M+H]^{3+}$ molecular ion marked by the asterisk in (A) was selected for ESI-MS/MS. The fragmentation pattern of the disaccharide linked to the peptide is shown in the inset. (C) Averaged ESI-MS spectrum of the LC elution containing the MC peptide of SP(-)-FbpC-MC demonstrated an absence of glycosylation. (D) Averaged ESI-MS spectrum corresponding to the LC elution containing the MC peptide of SodA-MC. The region of the spectrum denoted by the asterisk is enlarged in the inset and reveals a minor $[M+H]^{2+}$ molecular ion (1156.4 m/z) corresponding to the MC with a single hexose residue. (E) Averaged ESI-MS spectrum corresponding to the LC elution containing the MC peptide of ESAT-6-MC. The region of the spectrum denoted by the asterisk is enlarged in the inset and reveals a minor $[M+H]^{3+}$ molecular ion (1312.5 m/z) for the MC with a single hexose residue. (F) Averaged ESI-MS spectrum corresponding to the LC elution containing the MC peptide of SP(+)-ESAT-6-MC demonstrated a substantial level of glycosylation. (G) The dominant $[M+H]^{3+}$ 1312.2 m/z molecular ion marked by an asterisk in (F) was selected for ESI-MS/MS. The fragmentation pattern of the monosaccharide linked to the peptide is shown in the inset.

and lack SPs (19). To test the hypothesis that only Sec-translocation supports protein glycosylation, three additional recombinant fusion proteins were produced and assessed for glycosylation. The first two (SodA-MC and ESAT-6-MC) consisted of full-length SodA and ESAT-6 with C-terminal fusions to the Apa MC and a hexa-His tag (Fig. 2A). The final fusion SP(+)ESAT-6-MC had the SP and 34 additional amino acids of FbpC fused to the N terminus of ESAT-6-MC (Fig. 2A). This later chimera was designed to alter the natural export of ESAT-6 such that it was directed to the Sec-translocase. To further assure Sec-translocation of SP(+)ESAT-6-MC, a *Mtb* Δ RD1 mutant lacking ESAT-6 and its corresponding secretion machinery was used for recombinant expression (20). The SodA-MC and ESAT-6-MC fusions were abundant in all subcellular fractions of *Mtb*, including the CF and CW (Fig. 2B). SP(+)ESAT-6-MC, however, localized primarily to the cytoplasmic membrane and CW fractions of *Mtb* and *Mtb* Δ RD1 (Fig. 2B).

LC-ESI-MS/MS analyses of SodA-MC and ESAT-6-MC purified from the CF or CW of *Mtb* demonstrated low levels of glycosylation. Dominant molecular ions of $[M+H]^{2+}$ 1075.5 *m/z* and $[M+H]^{3+}$ 1258.4 *m/z* were observed and corresponded to the nonglycosylated MC peptides of SodA-MC and ESAT-6-MC (Fig. 3, D and E). Also observed were molecular ions (1156.4 *m/z* and 1312.5 *m/z*) for doubly and triply charged MC of SodA-MC and ESAT-6-MC that had single hexose residues. These molecular ions, however, were barely detectable compared with those of the nonglycosylated MCs (Fig. 3, D and E). The scarcity of glycosylation on SodA-MC and ESAT-6-MC was confirmed by the lack of ConA reactivity to purified SodA-MC and ESAT-6-MC (fig S3). The Sec-translocase-directed SP(+)ESAT-6-MC purified from the CW of *Mtb* and *Mtb* Δ RD1 was subjected to LC-ESI-MS/MS and produced two major $[M+H]^{3+}$ molecular ions (1258.4 and 1312.2 *m/z*) corresponding to the MC with zero and one hexose residue (Fig. 3F). Moreover, and unlike ESAT-6-MC, the 1312.2 *m/z* molecular ion was clearly dominant, demonstrating relatively abundant glycosylation (compare Fig. 3, E and F), and this glycosylation was confirmed by MS/MS fragmentation of the 1312.2 *m/z* molecular ion (Fig. 3G).

These data support the hypothesis that, analogous to eukaryotic systems, Sec-translocation is required for protein mannosylation in *Mtb*. Secondary structures surrounding sites of *O*-glycosylation are important for glycosyltransferase recognition (5), and in prokaryotes only proteins maintained in an unfolded state are targeted for Sec-dependent translocation (21). Therefore, the requirement of Sec-translocation for protein *O*-mannosylation in *Mtb* may be associated with substrate structural requirements and the export of unfolded proteins. Protein folding before translocation may also explain the

lack of glycosylation on the SodA and ESAT-6 fusion products exported by Sec-independent pathways (22, 23). Alternatively, efficient *O*-glycosylation of Sec-exported proteins could result from physical interactions between the Sec complex and a protein mannosyltransferase of *Mtb*. This would be similar to the proximity (30 to 40 Å) of the Sec61 translocon pore and the oligosaccharyltransferase (OST) complex and PMTs responsible for *N*- and *O*-glycosylation in eukaryotes (24, 25).

Our current data, along with the recent discovery of a eukaryotic-like OST *N*-glycosylation system in *Campylobacter jejuni* (13), demonstrate a previously unrealized evolutionary conservation between protein glycosylation pathways of eukaryotes and those of specific bacterial species. Moreover, the analogies to protein mannosylation of eukaryotes enable the establishment of a proposed mannoprotein assembly pathway in mycobacteria (fig. S4).

References and Notes

1. R. K. Upreti, M. Kumar, V. Shankar, *Proteomics* **3**, 363 (2003).
2. M. A. Schmidt, L. W. Riley, I. Benz, *Trends Microbiol.* **11**, 554 (2003).
3. K. M. Dobos, K. H. Khoo, K. M. Swiderek, P. J. Brennan, J. T. Belisle, *J. Bacteriol.* **178**, 2498 (1996).
4. S. L. Michell et al., *J. Biol. Chem.* **278**, 16423 (2003).
5. S. Strahl-Bolsinger, M. Gentzsch, W. Tanner, *Biochim. Biophys. Acta* **1426**, 297 (1999).
6. T. Garbe et al., *Infect. Immun.* **61**, 260 (1993).
7. C. Espitia, R. Mancilla, *Clin. Exp. Immunol.* **77**, 378 (1989).
8. F. Romain et al., *Infect. Immun.* **67**, 5567 (1999).
9. C. Horn et al., *J. Biol. Chem.* **274**, 32023 (1999).
10. H. D. Brightbill et al., *Science* **285**, 732 (1999).

11. B. A. Bensing, B. W. Gibson, P. M. Sullam, *J. Bacteriol.* **186**, 638 (2004).
12. C. Josenhans, L. Vossebein, S. Friedrich, S. Suerbaum, *FEMS Microbiol. Lett.* **210**, 165 (2002).
13. M. Wacker et al., *Science* **298**, 1790 (2002).
14. H. N. Cooper et al., *Glycobiology* **12**, 427 (2002).
15. V. Gierbach, T. Zeller, M. Priesmeier, S. Strahl-Bolsinger, *J. Biol. Chem.* **275**, 19288 (2000).
16. T. Willer, M. C. Valero, W. Tanner, J. Cruces, S. Strahl, *Curr. Opin. Struct. Biol.* **13**, 621 (2003).
17. J. T. Belisle et al., *Science* **276**, 1420 (1997).
18. D. R. Ronning et al., *Nat. Struct. Biol.* **7**, 141 (2000).
19. P. Andersen, D. Askgaard, L. Ljungqvist, J. Bennedsen, I. Heron, *Infect. Immun.* **59**, 1905 (1991).
20. T. Hsu et al., *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12420 (2003).
21. P. N. Danese, T. J. Silhavy, *Annu. Rev. Genet.* **32**, 59 (1998).
22. S. A. Stanley, S. Raghavan, W. W. Hwang, J. S. Cox, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13001 (2003).
23. M. Braunstein, B. J. Espinosa, J. Chan, J. T. Belisle, W. R. Jacobs Jr., *Mol. Microbiol.* **48**, 453 (2003).
24. I. M. Nilsson, G. von Heijne, *J. Biol. Chem.* **268**, 5798 (1993).
25. M. Ecker et al., *EMBO Rep.* **4**, 628 (2003).
26. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
27. We thank P. J. Hill for technical assistance and W. R. Jacobs Jr. (Howard Hughes Medical Institute, Albert Einstein College of Medicine) for the *M. tuberculosis* Δ RD1 mutant. This work was supported by NIH National Institute of Allergy and Infectious Diseases grant R01 AI-44042.

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5736/941/DC1

Materials and Methods

Figs. S1 to S4

References

3 May 2005; accepted 17 June 2005
10.1126/science.1114347

Regulation of Blood Glucose by Hypothalamic Pyruvate Metabolism

Tony K. T. Lam, Roger Gutierrez-Juarez, Alessandro Pocai, Luciano Rossetti*

The brain keenly depends on glucose for energy, and mammals have redundant systems to control glucose production. An increase in circulating glucose inhibits glucose production in the liver, but this negative feedback is impaired in type 2 diabetes. Here we report that a primary increase in hypothalamic glucose levels lowers blood glucose through inhibition of glucose production in rats. The effect of glucose requires its conversion to lactate followed by stimulation of pyruvate metabolism, which leads to activation of adenosine triphosphate (ATP)-sensitive potassium channels. Thus, interventions designed to enhance the hypothalamic sensing of glucose may improve glucose homeostasis in diabetes.

Diabetic hyperglycemia is due in part to an inappropriately elevated rate of liver glucose production (1). This is paradoxical, since hyperglycemia should restrain glucose production (2–4). Hypothalamic glucose sensing plays an important role in preserving nutrient homeostasis (5–10). Here we test the hypothesis that activation of neuronal pyruvate flux is required for hypothalamic glucose sensing and for control of blood glucose levels and liver

glucose metabolism. Blood glucose gains access to the CNS by means of a facilitated transport system (11), and indeed, the extracellular concentration of glucose in the brain is significantly lower (~2 mM) than its concentration in blood (~5 mM). However, it remains controversial whether neurons directly metabolize glucose to pyruvate or use lactate derived from the anaerobic glycolysis of glucose within astrocytes (12–14) (Fig. 1A).

The transfer of glucose-derived lactate from astrocytes to neurons is referred to as the "astrocyte-neuron lactate shuttle" (12, 15).

To examine the central effects of glucose on systemic glucose homeostasis, we infused D-glucose (2 mM) into the third cerebral ventricle of conscious rats. Intracerebro-ventricular (ICV) glucose resulted in a 69% increase in the hypothalamic glucose concentration (from 0.88 to 1.48 $\mu\text{mol/g}$) and in decreased blood glucose and insulin levels (Fig. 1B) (table S1). In the presence of basal insulin (table S1), the rate of glucose infusion required to maintain euglycemia was marginal in rats receiving ICV mannitol (2 mM). However, when rats received ICV glucose, glucose had to be infused systemically to prevent hypoglycemia (Fig. 1C). The increased glucose infusion was due to suppression of liver glucose production (Fig. 1D).

Glucose production represents the net contribution of gluconeogenesis and glycogenolysis. However, a portion of glucose entering the liver

by means of phosphorylation of glucose is also a substrate for dephosphorylation by means of glucose-6-phosphatase (G6Pase, encoded by *G6pc*), creating a futile cycle. To delineate the mechanisms by which central activation of glucose metabolism modulates blood glucose and liver glucose homeostasis, we estimated the in vivo flux through G6Pase and the relative contribution of gluconeogenesis and glycogenolysis to glucose output. ICV glucose markedly decreased the flux through G6Pase and the hepatic expression of *G6pc* (Fig. 1E; fig. S1A). Inhibition in both gluconeogenesis and glycogenolysis accounted for the decrease in glucose production (fig. S1A). Conversely, the levels of glucoregulatory hormones, the rate of glucose utilization, and the hepatic expression of the gluconeogenic gene phosphoenolpyruvate carboxykinase, encoded by *Pck1*, were unchanged (fig. S1A, table S1). These results indicate that the potent effects of central glucose on liver glucose fluxes are largely due to marked inhibition of hepatic G6Pase.

According to the astrocyte-neuron lactate shuttle hypothesis (12), central administration of glucose increases the formation of lactate in astrocytes (with the conversion of each glucose molecule into two molecules of lactate) to provide extracellular lactate for neurons, which

in turn rapidly convert it to pyruvate (Fig. 1A). Therefore, we investigated whether ICV lactate could recapitulate the effects of ICV glucose on blood glucose and liver glucose metabolism. Indeed, the ICV infusion of L-lactate (5 mM) decreased blood glucose and insulin levels (Fig. 1B; table S2) compared with equimolar infusions of its isomer D-lactate, which cannot be metabolized. In the presence of basal levels of circulating insulin (table S2), ICV L-lactate increased the rate of glucose infusion required to maintain euglycemia (Fig. 1C). This increase was due to suppression of liver glucose production (Fig. 1D). The flux through G6Pase, the hepatic expression of *G6pc*, gluconeogenesis, and glycogenolysis were all decreased in rats receiving ICV L-lactate (Fig. 1E; fig. S1B). However, the levels of glucoregulatory hormones, the rate of glucose utilization, and the hepatic expression of *Pck1* were unchanged (fig. S1B, table S2). Thus, a moderate increase in the central availability of either glucose or lactate is sufficient to lower blood glucose by means of rapid changes in liver glucose metabolism and gene expression.

The use of extracellular lactate by neurons requires its conversion to pyruvate by the enzyme lactic dehydrogenase (mainly *LDH-B* in neurons) (16). If the effects of ICV lactate on

Departments of Medicine and Molecular Pharmacology, Diabetes Research Center, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA.

*To whom correspondence should be addressed. E-mail: rossetti@aecom.yu.edu

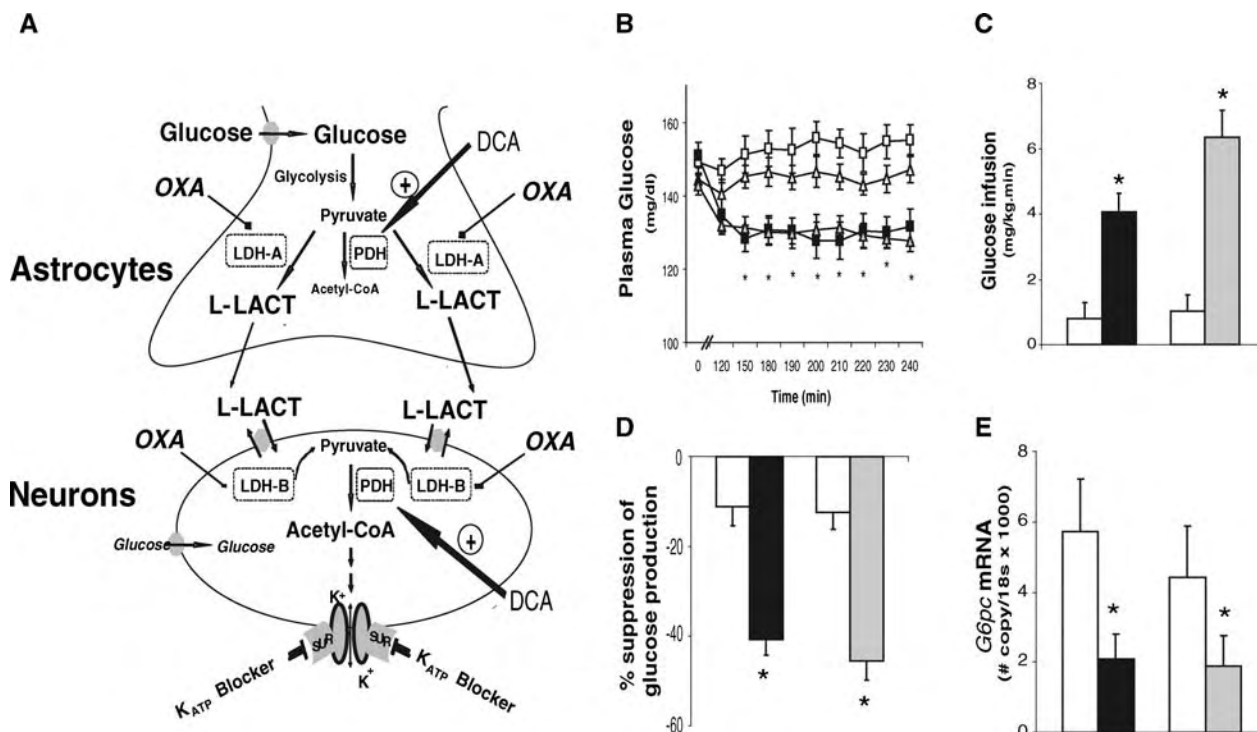


Fig. 1. Central administration of glucose or lactate lowers blood glucose by means of suppression of endogenous glucose production. (A) Schematic representation of hypothesis. Glucose enters astrocytes where it is metabolized to pyruvate by means of glycolysis. Pyruvate is preferentially converted to L-lactate (L-LACT) by lactate dehydrogenase (LDH-A in astrocytes) and can be taken up by neurons to generate pyruvate by means of LDH (LDH-B in neurons). Finally, pyruvate is converted to acetyl-CoA by pyruvate dehydrogenase (PDH). The increased flux through PDH ultimately leads to the activation of K_{ATP} channels. Oxamate (OXA) is a

competitive inhibitor of LDH. Dichloroacetate (DCA) is an activator of PDH. (B) Effect of the central administration of glucose (black squares and bars; $n = 6$) or L-lactate (gray triangles and bars; $n = 6$), mannitol (open squares, white bars; $n = 5$), or D-lactate (open triangles, white bars; $n = 5$) on glucose levels. Glucose and L-lactate lowered blood glucose, whereas mannitol or D-lactate did not. During clamp, (C) glucose and L-lactate increased glucose infusion, (D) decreased glucose production, and (E) decreased liver *G6pc* mRNA relative to mannitol or D-lactate. * $P < 0.01$ versus mannitol or D-lactate. Values are means \pm SEM.

glucose homeostasis are mediated by neuronal lactate use, the inhibition of LDH should negate these effects. Therefore, we examined whether impeding the conversion of lactate to pyruvate negates the metabolic effects of the

central administration of L-lactate (Fig. 1A). Oxamate is a competitive inhibitor of LDH (17). The coinfusion of oxamate (50 mM) abolished the effects of ICV L-lactate on blood glucose and insulin levels (Fig. 2A; table S2).

During clamp, ICV oxamate also negated the effects of ICV L-lactate on glucose infusion (Fig. 2B), glucose production (Fig. 2C), G6Pase flux (fig. S2A), *G6pc* mRNA (Fig. 2D), gluconeogenesis (Fig. 2E), and glycogen-

Fig. 2. Central inhibition of LDH negates the effects of L-lactate and glucose on glucose homeostasis. (A) The ICV coinfusion of the LDH inhibitor oxamate with glucose (black squares, dotted lines; $n = 6$) or with L-lactate (gray triangles, dotted lines; $n = 6$) abolished the blood glucose-lowering effect of glucose (black squares; $n = 6$) or L-lactate (gray triangles; $n = 6$). Oxamate alone (white squares) did not modify blood glucose. $*P < 0.01$ versus oxamate and versus oxamate with glucose or L-lactate. (B) During clamp, ICV coinfusion of oxamate with L-lactate (right gray bars) or glucose (right black bars) negated the effects of ICV L-lactate (left gray bars) or glucose (left black bars) on the rate of glucose infusion, (C) on the suppression of glucose production, (D) on liver *G6pc* mRNA, (E) on gluconeogenesis, and (F) on glycogenolysis. $*P < 0.05$ versus oxamate and versus oxamate with glucose or L-lactate. Values are means \pm SEM.

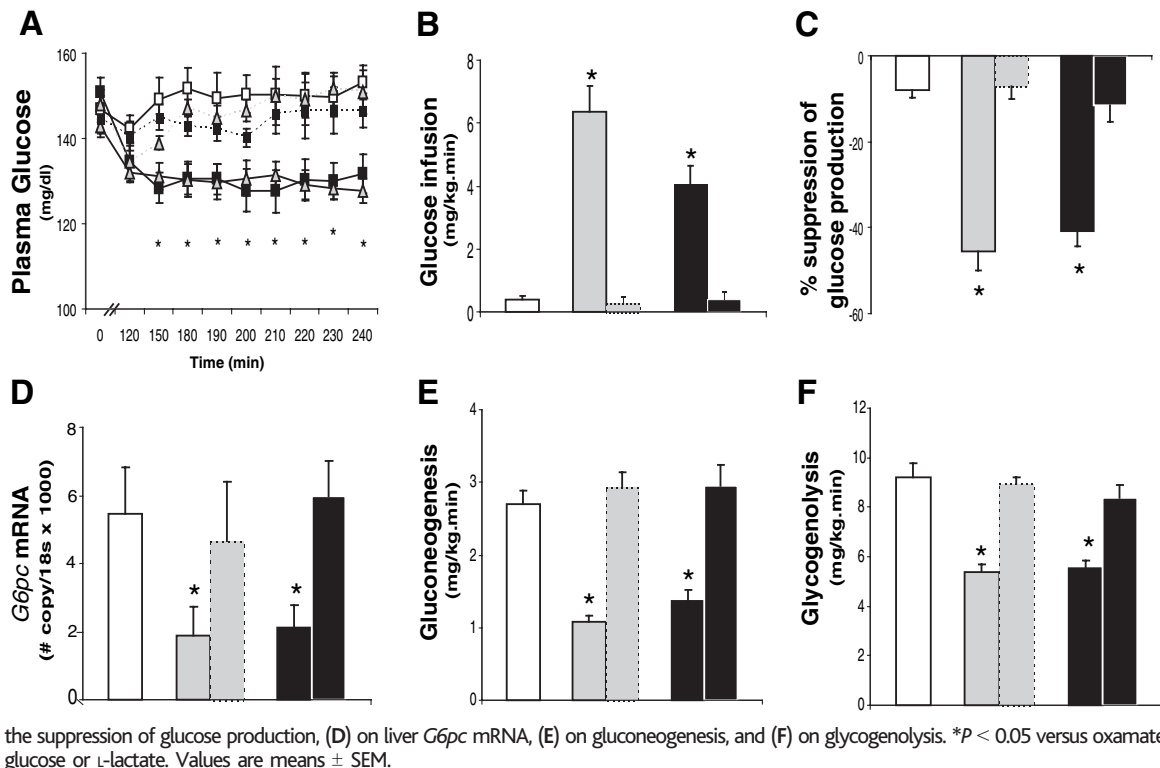
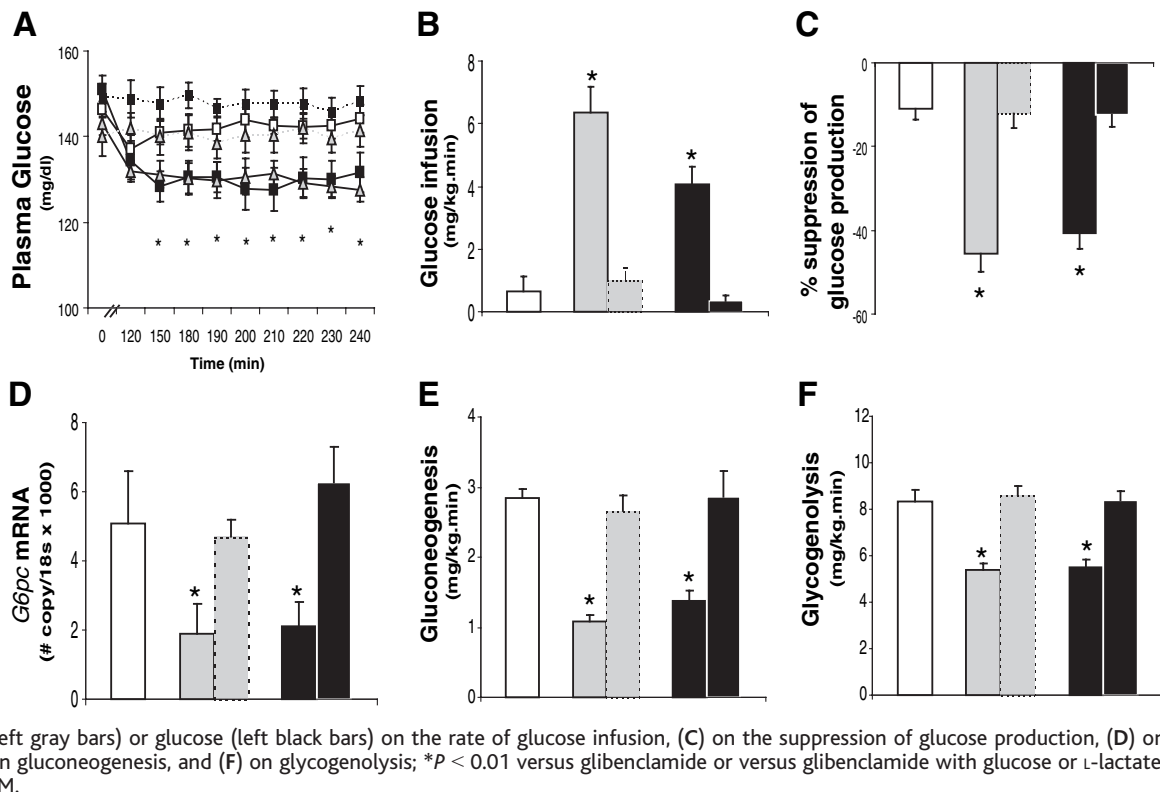


Fig. 3. Central administration of a K_{ATP} channels blocker negates the effects of L-lactate and glucose on glucose homeostasis. (A) The ICV coinfusion of the K_{ATP} channels blocker glibenclamide with glucose (black squares, dotted lines; $n = 6$) or L-lactate (gray triangles, dotted lines; $n = 6$) abolished the blood glucose lowering effect of glucose (black squares; $n = 6$) or L-lactate (gray triangles; $n = 6$). Glibenclamide alone (white squares) did not modify blood glucose; $*P < 0.01$ versus glibenclamide and versus glibenclamide with glucose or L-lactate. (B) During clamp, ICV coinfusion of glibenclamide with L-lactate (right gray bars) or glucose (right black bars) negated the effects of ICV L-lactate (left gray bars) or glucose (left black bars) on the rate of glucose infusion, (C) on the suppression of glucose production, (D) on liver *G6pc* mRNA, (E) on gluconeogenesis, and (F) on glycogenolysis; $*P < 0.01$ versus glibenclamide or versus glibenclamide with glucose or L-lactate. Values are means \pm SEM.



olysis (Fig. 2F). These data indicate that the metabolism of lactate to pyruvate is an obligatory biochemical step for the regulation of liver glucose homeostasis by central lactate.

Glucose could be directly metabolized to pyruvate within neurons or it may first be converted to lactate, the metabolism of which then generates pyruvate. If the conversion of glucose to lactate is required for the effects of ICV glucose on liver glucose homeostasis, the coinfusion of oxamate should also prevent the metabolic effects of ICV glucose. Indeed, the central administration of oxamate abolished the effects of ICV glucose on blood glucose and insulin levels (Fig. 2A; table S1). During

clamp, ICV oxamate also negated the effects of ICV D-glucose on glucose infusion (Fig. 2B), glucose production (Fig. 2C), G6Pase flux (fig. S2B), *G6pc* mRNA (Fig. 2D), gluconeogenesis (Fig. 2E), and glycogenolysis (Fig. 2F). Together these findings reveal a negative-feedback loop between the central availability of glucose or lactate and the regulation of liver glucose homeostasis.

The activation of hypothalamic adenosine triphosphate (ATP)-sensitive potassium (K_{ATP}) channels is critical for modulation of blood glucose levels and liver glucose fluxes (18), as well as for the metabolic effects of other nutrient-dependent signals (18–20). To

examine the role of central K_{ATP} channels in carbohydrate sensing, we coinjected ICV the K_{ATP} blocker glibenclamide (100 μ M) with lactate or glucose (Fig. 1A). This inhibitor negated the effects of both ICV lactate and glucose on blood glucose (Fig. 3A). Furthermore, during clamp, the central infusion of glibenclamide prevented the increase in the rate of glucose infusion (Fig. 3B), as well as the decrease in glucose production (Fig. 3C) induced by either ICV lactate or glucose. Similarly, the K_{ATP} blocker abolished the effects of central lactate and glucose on G6Pase flux (fig. S3, A and B), *G6pc* mRNA (Fig. 3D), gluconeogenesis (Fig. 3E), and glycogenolysis (Fig.

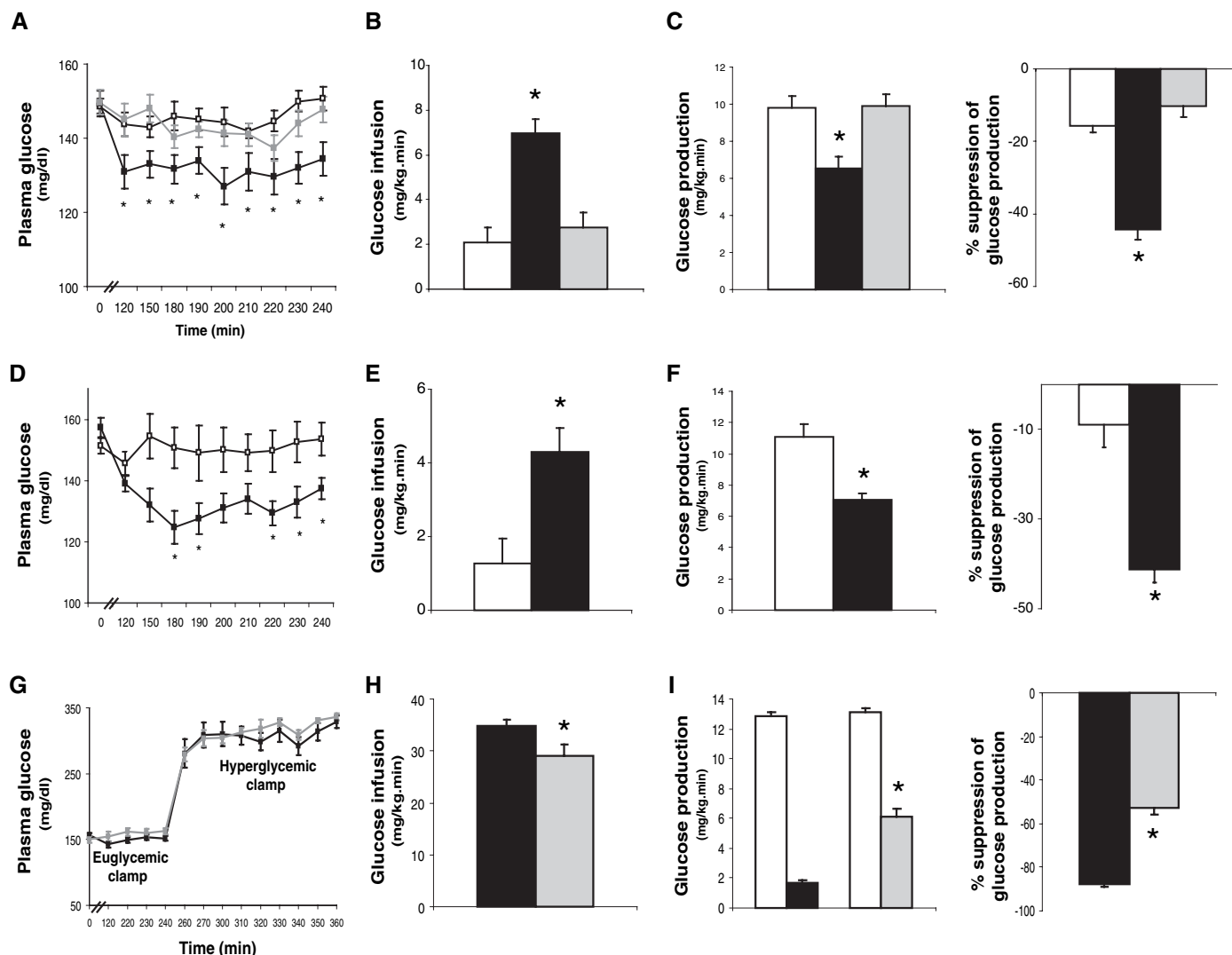


Fig. 4. Intrahypothalamic infusions of glucose, PDH activator, or LDH inhibitor regulate glucose production. (A) Bilateral cannulae were placed within the parenchyma of the mediobasal hypothalamus. The IH coinfusion of the K_{ATP} channels blocker glibenclamide with glucose (gray squares; $n = 5$) abolished the blood glucose-lowering effect of glucose (black squares; $n = 5$). Glibenclamide alone (white squares) did not modify blood glucose. $*P < 0.01$ versus glibenclamide or glucose + glibenclamide. (B) During clamp, IH coinfusion of glibenclamide with glucose (gray bars) negated the effects of IH glucose (black bars) on the rate of glucose infusion, and (C) on glucose production; $*P < 0.01$ versus glibenclamide or glucose + glibenclamide; (D) The IH infusion of the PDH activator

dichloroacetate (DCA, black bars; $n = 5$) lowered blood glucose, (E) increased glucose infusion, and (F) decreased glucose production; $*P < 0.01$ versus vehicle (white bars). (G) During pancreatic-basal insulin clamps, circulating blood glucose was doubled and then maintained at these levels in the presence (gray bars, $n = 4$) or absence (black bars, $n = 4$) of IH oxamate. IH oxamate resulted in (H) decreased rates of glucose infusion during pancreatic-hyperglycemic clamps, (I) did not alter glucose production in the presence of euglycemia (white bar), but increased glucose production during the hyperglycemia so that the suppressive effect of hyperglycemia on glucose production was diminished. $*P < 0.001$, IH OXA versus IH vehicle. Values are means \pm SEM.

3F). Glucoregulatory hormone levels, glucose use, and liver *Pck1* mRNA were unchanged (fig. S3, A and B; tables S1 and S2). These findings indicate that the central activation of K_{ATP} channels bridges the hypothalamic sensing of glucose and lactate with the regulation of liver glucose homeostasis.

To investigate the anatomical localization of these effects, we infused a dose of glucose that was lower (by a factor of 15) than that used in the ICV studies bilaterally within the parenchyma of the mediobasal hypothalamus (intra-hypothalamic, IH) (fig. S4, A and B). IH glucose (2 mM) lowered circulating glucose levels (Fig. 4A). During clamp, glucose had to be infused systemically to prevent hypoglycemia when glucose was administered IH (Fig. 4B). The increased glucose infusion was due to a reduction in liver glucose production (Fig. 4C); however, the rate of glucose use was not affected (fig. S4C). Importantly, the IH coinfusion of glibenclamide negated the effects of IH glucose on blood glucose, glucose infusion, and glucose production (Fig. 4, A to C). Thus, the increased availability of glucose leads to the activation of a K_{ATP} -dependent pathway within the mediobasal hypothalamus that is sufficient to lower blood glucose by means of the inhibition of glucose production. Because the relevant (Sur1-containing) K_{ATP} channels do not appear to be expressed in glial cells (18, 19), it is likely that this step in hypothalamic glucose sensing is occurring in neurons.

The entry of pyruvate in the tricarboxylic cycle (TCA) is governed by its conversion to acetyl-coenzyme A (acetyl-CoA) by pyruvate dehydrogenase (PDH). If the flux of pyruvate into the TCA cycle is the key signal generated by the central administration of glucose or lactate, then the stimulation of pyruvate flux should recapitulate their effects on glucose homeostasis (Fig. 1A; fig. S4A). Dichloroacetate (DCA) activates PDH by means of inhibition of PDH kinase and increases pyruvate flux in neurons and in astrocytes (21). The delivery of DCA (1 mM) within the mediobasal hypothalamus lowered circulating glucose levels (Fig. 4D). During clamp, the IH administration of DCA increased the rate of glucose infusion required to maintain euglycemia (Fig. 4E). This effect was due to suppression of liver glucose production (Fig. 4F), rather than to increased glucose utilization (fig. S4D). Because activation of PDH in astrocytes would actually decrease lactate formation, our results with DCA are consistent with neuronal activation of PDH in vivo. This in turn reproduced the metabolic effects of central glucose or lactate. Because the enhanced conversion of pyruvate to acetyl-CoA should lead to increased flux into the TCA cycle (Fig. 1A), it seems plausible that the neuronal TCA cycle serves as a biochemical sensor for carbohydrate availability in the hypothalamus, which in turn regulates liver glucose homeostasis.

To estimate the contribution of hypothalamic glucose sensing to the overall effects of circulating glucose on liver glucose homeostasis, we used the pancreatic-hyperglycemic clamp technique to double the circulating glucose levels in rats receiving IH oxamate or IH vehicle (Fig. 4G). This level of hyperglycemia led to a twofold increase in the hypothalamic glucose levels (to 2.1 μ mol/g). The rate of glucose infusion required to maintain hyperglycemia was less with IH oxamate than with IH vehicle (Fig. 4H). Hyperglycemia inhibited glucose production by 88% in the presence of IH vehicle and by only 53% in the presence of IH oxamate (Fig. 4I). Thus, blocking the metabolism of lactate in the hypothalamus results in a loss of 40% of the inhibitory action of circulating glucose on glucose production, which indicates that the neuronal circuit engaged in response to increased hypothalamic glucose metabolism plays an important role in restraining glucose production in response to an increase in the circulating glucose levels.

Our results are also relevant to the ongoing debate on the physiological role of the astrocyte-neuron lactate shuttle (12). According to this hypothesis, an increase in neuronal activity raises the extracellular levels of the excitatory neurotransmitter glutamate, which in turn activates glycolysis and lactate production in astrocytes. Consistent with this notion, the IH infusion of glutamate (1 mM) recapitulated the effects of IH glucose on blood glucose levels and on liver glucose production (22). Because the entry of pyruvate into the TCA cycle is a dominant pathway in neuronal energetics, the role of pyruvate flux in glucose sensing appears to link neuronal activity within this hypothalamic region to the regulation of liver glucose homeostasis. Alterations in the function of mitochondria have been suggested to play important roles in the etiology of metabolic changes associated with aging, diabetes, and obesity (23–25). These alterations, if extended to hypothalamic neurons, could result in decreased flux through PDH and, therefore, could hamper the responses to nutritional cues converging on the increased availability of pyruvate.

The arcuate nucleus of the hypothalamus is emerging as a major site for the integration of multiple nutritional (19, 26) and hormonal (18, 20, 27–31) signals, which are central to the modulation of liver glucose homeostasis. Type 2 diabetes and the metabolic syndrome are typical examples of diseases the prevalence of which is dependent on environmental and/or nutritional factors operating on genetic susceptibility. Impairment in the biochemical sensing of carbohydrates may represent a basic underpinning for defects in the regulation of food intake (6), β -cell function (32), and liver glucose homeostasis (4).

In conclusion, neurons appear to have developed mechanisms for glucose sensing designed to protect them from hypoglycemic injury by

triggering the rapid secretion of counterregulatory hormones in response to low extracellular glucose levels (8–10). Here, we have shown that moderate increases in extracellular glucose levels within a specific region of the hypothalamus are sufficient to lower blood glucose levels through a robust inhibition of liver glucose production. This discovery paves the way for biochemical interventions designed to restore central glucose sensing and glucose homeostasis.

References and Notes

- S. I. Taylor, *Cell* **97**, 9 (1999).
- L. Sacca, R. Hendler, R. S. Sherwin, *J. Clin. Endocrinol. Metab.* **47**, 1160 (1978).
- L. Rossetti et al., *J. Clin. Invest.* **92**, 1126 (1993).
- M. Mevorach et al., *J. Clin. Invest.* **102**, 744 (1998).
- J. D. Davis, D. Wirtshafter, K. E. Asin, D. Brief, *Science* **212**, 81 (1981).
- S. C. Woods, L. D. McKay, *Science* **202**, 1209 (1978).
- D. A. Thompson, R. G. Campbell, *Science* **198**, 1065 (1977).
- M. A. Borg, R. S. Sherwin, W. P. Borg, W. V. Tamborlane, G. I. Shulman, *J. Clin. Invest.* **99**, 361 (1997).
- B. E. Levin, V. H. Routh, L. Kang, N. M. Sanders, A. A. Dunn-Meynell, *Diabetes* **53**, 2521 (2004).
- P. E. Cryer, S. N. Davis, H. Shamoon, *Diabetes Care* **26**, 1902 (2003).
- F. Maher, S. J. Vannucci, I. A. Simpson, *FASEB J.* **8**, 1003 (1994).
- L. Pellerin, P. J. Magistretti, *Science* **305**, 50 (2004).
- L. Pellerin, P. J. Magistretti, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10625 (1994).
- P. J. Magistretti, L. Pellerin, D. L. Rothman, R. G. Shulman, *Science* **283**, 496 (1999).
- K. A. Kasischke, H. D. Vishwasrao, P. J. Fisher, W. R. Zipfel, W. W. Webb, *Science* **305**, 99 (2004).
- P. G. Bittar, Y. Charnay, L. Pellerin, C. Bouras, P. J. Magistretti, *J. Cereb. Blood Flow Metab.* **16**, 1079 (1996).
- G. A. Brooks, H. Dubouchaud, M. Brown, J. P. Sicurello, C. E. Butz, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1129 (1999).
- A. Pocai et al., *Nature* **434**, 1026 (2005).
- T. K. Lam et al., *Nat. Med.* **11**, 320 (2005).
- S. Obici, B. B. Zhang, G. Karkianis, L. Rossetti, *Nat. Med.* **8**, 1376 (2002).
- Y. Itoh et al., *Proc. Natl. Acad. Sci. U.S.A.* **100**, 4879 (2003).
- R. Gutierrez-Juarez, T. K. T. Lam, A. Pocai, L. Rossetti, unpublished observations.
- B. B. Lowell, G. I. Shulman, *Science* **307**, 384 (2005).
- K. F. Petersen, S. Dufour, D. Befroy, R. Garcia, G. I. Shulman, *N. Engl. J. Med.* **350**, 664 (2004).
- K. F. Petersen et al., *Science* **300**, 1140 (2003).
- S. Obici, Z. Feng, A. Arduini, R. Conti, L. Rossetti, *Nat. Med.* **9**, 756 (2003).
- M. W. Schwartz, D. Porte Jr., *Science* **307**, 375 (2005).
- R. Coppari et al., *Cell Metab.* **1**, 63 (2005).
- J. S. Flier, *Cell* **116**, 337 (2004).
- G. J. Morton et al., *Endocrinology* **144**, 2016 (2003).
- P. Cohen et al., *J. Clin. Invest.* **108**, 1113 (2001).
- L. Rossetti, G. I. Shulman, W. Zawulich, R. A. DeFronzo, *J. Clin. Invest.* **80**, 1037 (1987).
- The study protocol was approved by the Institutional Animal Care and Use Committee of the Albert Einstein College of Medicine. We thank H. Zhang, B. Liu, and S. Gaweda for technical assistance. Supported by grants from the NIH to L.R. (DK 45024, DK 48321 and AG 21654) and to the Albert Einstein College of Medicine Diabetes Research and Training Center (DK 20541). T.K.T.L. is supported by a Training Grant from the National Institute on Aging (T32-AG023475).

Supporting Online Materials

www.sciencemag.org/cgi/content/full/309/5736/943/DC1

Materials and Methods

Fig. S1 to S4

Table S1 and S2

References and Notes

10 March 2005; accepted 21 June 2005

10.1126/science.1112085

Hemodynamic Signals Correlate Tightly with Synchronized Gamma Oscillations

Jörn Niessing,¹ Boris Ebisch,¹ Kerstin E. Schmidt,¹
Michael Niessing,¹ Wolf Singer,¹ Ralf A. W. Galuske^{1,2*}

Functional imaging methods monitor neural activity by measuring hemodynamic signals. These are more closely related to local field potentials (LFPs) than to action potentials. We simultaneously recorded electrical and hemodynamic responses in the cat visual cortex. Increasing stimulus strength enhanced spiking activity, high-frequency LFP oscillations, and hemodynamic responses. With constant stimulus intensity, the hemodynamic response fluctuated; these fluctuations were only loosely related to action potential frequency but tightly correlated to the power of LFP oscillations in the gamma range. These oscillations increase with the synchrony of synaptic events, which suggests a close correlation between hemodynamic responses and neuronal synchronization.

Blood oxygenation level-dependent (BOLD) imaging methods are powerful tools to investigate brain function, and they are particularly well suited for studies on the human brain (1, 2). However, there is still debate about which aspects of neuronal activity are reflected by the amplitude of hemodynamic responses. BOLD responses are positively correlated with action potentials (3, 4) and the amplitude of evoked potentials (5–7). The latter depends not only on the number and discharge rates of activated neurons but also on the temporal coherence of action potentials and synaptic activity (8). The BOLD-response may thus not only be influenced by the amplitude but also by the temporal structure of neuronal discharges. This possibility is supported by recent studies, which have shown that LFP oscillations correlate better with BOLD signals than with discharge rates (9, 10) and that synaptic activity elevates BOLD signals even in the absence of action potentials (11).

LFP-oscillations and, thus, the temporal coherence of synaptic activity vary substantially in amplitude and frequency as a function of stimulus configurations and central states. During sleep and relaxed wakefulness, oscillations occur at low frequencies (12), whereas high-frequency oscillations prevail during states of arousal and focused attention (13, 14). We studied the relations between hemodynamic signals, action potentials, and LFP oscillations in different frequency bands. Anesthetized adult cats were visually stimulated with moving whole-field gratings of different orientation, and electrical and hemodynamic activity in primary visual cortex was

monitored simultaneously with the use of implanted microelectrodes and optical imaging of intrinsic signals at 570 and 610 nm (1). Variations in the amplitude of cortical responses were induced by presenting visual stimuli at two contrast levels (27 and 97% for low contrast and high contrast, respectively). In a parallel approach, spontaneous response changes were examined while stimulus contrast was kept constant at 97% and the variability of oscillatory patterns of cortical activity was enhanced by electrical activation of the mesencephalic reticular formation (MRF). Stimulation of this structure has been shown to enhance oscillatory activity in the gamma range by means of activation of cholinergic afferents from the basal forebrain (15, 16). As a measure of the magnitude of hemodynamic responses, we used the signal amplitude in differential orientation maps calculated by subtraction of orientation maps from their orthogonal counterpart. The amplitude of electrical responses was assessed from the discharge frequency of multiunit activity (MUA) and the oscillatory patterning of responses from the power spectra of LFPs that were recorded from the same electrodes as the MUA (17).

Differential orientation maps resulting from stimulation with high-contrast gratings at 610 nm exhibited a markedly stronger signal amplitude than those obtained at low contrast. This is indicated by the enhanced contrast in the differential maps and the steeper gradients between regions responding to different orientations leading to a sharper delineation of orientation domains at high-contrast stimulation (Fig. 1A). Even though the time course of the hemodynamic responses was similar for both contrast levels, the peak amplitudes at low-contrast stimulation were on average ~31% lower ($P < 0.005$, Mann-Whitney U test) than at high-contrast stimulation (Fig.

1B). Simultaneously recorded MUA showed the typical phasic-tonic response pattern (Fig. 1C) and firing rates (averaged over the whole stimulus duration) were 28% lower with the low-contrast than with the high-contrast gratings ($P < 0.0001$, Mann-Whitney U test). The oscillatory LFP responses to both stimulus contrasts are shown in Fig. 1D. In these plots, power spectra from blocks of eight individual trials were averaged and then sorted in an ascending order according to the strength of the corresponding hemodynamic response. At both contrast levels oscillations occurred over a broad range of frequencies. However, on average, low-contrast stimuli evoked oscillatory responses at lower frequencies than did the high-contrast stimuli, which included predominantly oscillations in the upper gamma frequency range between 50 and 90 Hz (Fig. 1D). Interestingly, this relation also holds for responses evoked by stimuli of equal contrast. Neuronal responses in trial blocks associated with strong optical signals tended to oscillate at higher frequencies (Fig. 1D). Correlation analysis between hemodynamic signals and neuronal responses calculated across all trials and both contrast levels revealed a moderate correlation with spike rate ($r = 0.53$, Fig. 1E) and a somewhat stronger relation ($r = 0.6$, Fig. 1F) with the relative power of oscillations in the upper gamma band.

We also found a marked variability of hemodynamic responses even when contrast levels were kept constant. This variability occurred in a statelike manner in which phases of strong hemodynamic responses alternated with phases of weaker hemodynamic responsiveness. Thus, we kept the stimulus contrast constant at 97% and recorded electrophysiological and optical responses at 610 nm over several hours. For quantitative analysis, blocks of averaged results from eight repetitions of the eight different stimuli (17), which had been used for calculation of the hemodynamic responses, were grouped in three classes according to the strength of the respective optical response, and both optical and electrophysiological response variables were averaged within these groups. The peak amplitudes of the averaged hemodynamic responses differed significantly between the three groups. There were also marked differences in response duration; the larger responses decayed more slowly and lasted longer (Fig. 2A). Interestingly, the range of variability of hemodynamic response strength was as large as the range covered by responses evoked at different stimulus intensities (compare Figs. 1D and 2C). Neuronal firing rates also varied considerably between trials and these variations were related to the fluctuations of the hemodynamic responses (Fig. 2B). However, averaged firing rates differed only between the group with strong hemodynamic responses and the groups of medium and low response strength but not

¹Max Planck Institute for Brain Research, 60528 Frankfurt/M., Germany. ²Department of Biology, Technical University of Darmstadt, 64283 Darmstadt, Germany.

*To whom correspondence should be addressed. E-mail: galuske@mpih-frankfurt.mpg.de

between the latter two. By contrast, a clear difference between all three groups of hemodynamic responses existed with respect to the frequency distribution of the respective oscillatory responses in the LFPs. Low-frequency oscillations in the delta-, theta- and the alpha-frequency band (Fig. 2C) were most prominent in trial fraction 1, containing the trials with the weakest optical signal. The weakest high-frequency oscillations were observed in this fraction. With increasing hemodynamic response strength (trial fraction 2) oscillation frequency shifted from the theta and alpha band to the beta and lower gamma frequency band, and the strongest hemodynamic responses (trial fraction 3) were associated with the most prominent oscillations in the lower and upper gamma frequency band.

Figure 3A summarizes the relations between the strength of hemodynamic responses at 610 nm and the oscillation power in different LFP frequency bands for the entire data set. Low frequency activity in the delta band showed a strong negative relation with hemodynamic signal strength ($r_{\text{delta}} = -0.44$), whereas theta, alpha, and beta activities were more or less uncorrelated with this signal ($r_{\text{theta}} = -0.23$, $r_{\text{alpha}} = -0.03$, $r_{\text{beta}} = 0.18$). A weak positive relation existed for activity in the lower gamma band ($r_{\text{gamma1}} = 0.33$), and a strong positive correlation for oscillations in the upper gamma band ($r_{\text{gamma2}} = 0.64$). These relations were similar for trials without (Fig. 3A, black dots) and with MRF activation (Fig. 3A, gray dots). Optical signals measured at 570 nm exhibited a significant positive relationship exclusively with oscillations in the upper gamma band (Fig. 3B, $r_{\text{gamma2}} = 0.5$). In contrast, the correlation between firing rates and the optical signal at 610 nm was very low ($r = 0.19$) (Fig. 3C and fig. S2). At individual recording sites, firing rates usually varied with variations of the corresponding hemodynamic responses, but these correlations could be both positive and negative (r values ranging from -0.49 to 0.74). A consistent positive correlation was found for only 25% of the recording sites ($r > 0.35$). We also correlated firing rates with the power of gamma oscillations on a trial-by-trial basis and did not find a consistent relation (fig. S3).

The experiments with gratings of different contrast confirm the previous view that hemodynamic responses are positively correlated with stimulus intensity (18) and neuronal discharge rate (4, 9). Moreover, the analysis of LFP oscillations revealed that stronger stimuli shifted the frequency of LFP oscillations to higher values and that there is a particularly tight correlation between hemodynamic responses and LFP oscillations in the high gamma frequency range. The present findings closely link hemodynamic responses to the processes leading to LFP oscillations in the gamma frequency range. Both aspects that

contribute to the BOLD signal, blood volume changes (19) and changes in the deoxygenation status of hemoglobin (20), are positively correlated to the occurrence of gamma oscillations. Thus, our results are compatible with the finding that BOLD responses correlate better with the amplitude of LFP oscillations than with the discharge rate of cortical neurons (9, 10). Because LFPs predominantly reflect transmembrane currents associated with syn-

aptic activity (8), it had been proposed that BOLD responses reflected input rather than output activity (9, 11), a conclusion which has also been reached on the basis of theoretical assumptions on the neuronal energy budget (21). Our data go beyond this interpretation and suggest that the precision of synchronous firing of neurons within a cortical volume critically contributes to the magnitude of the hemodynamic response. The amplitude of

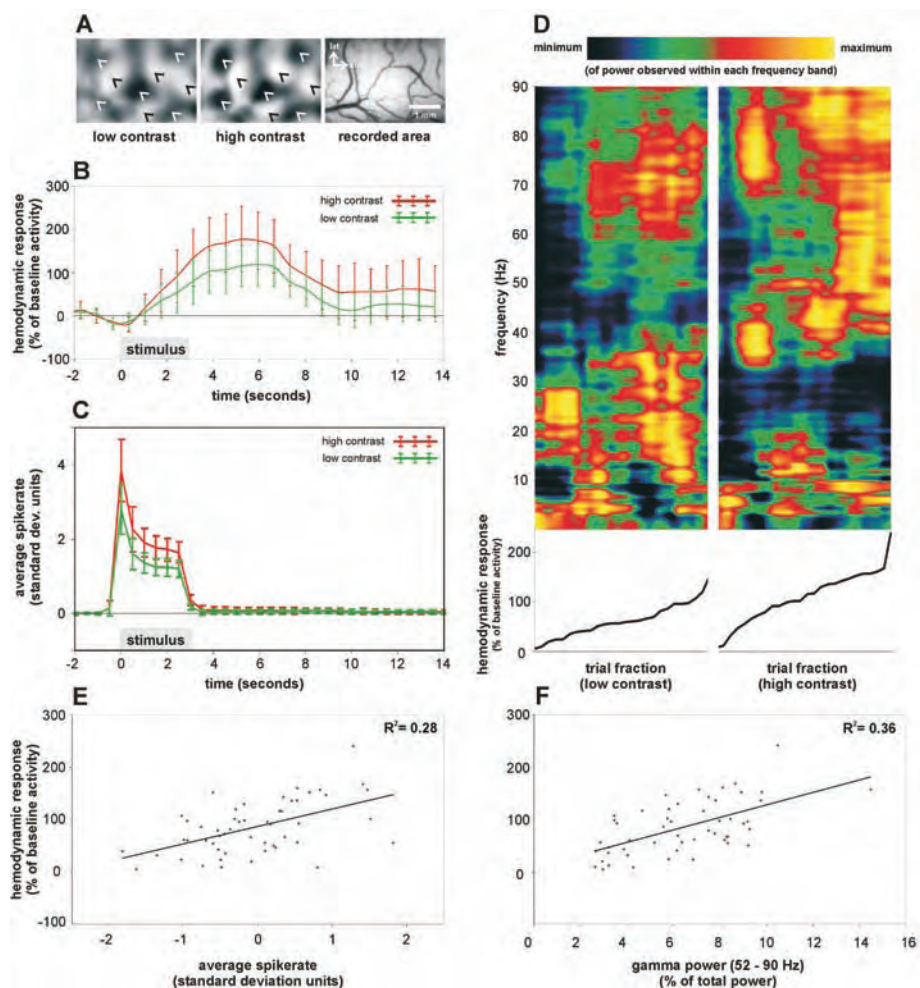


Fig. 1. Hemodynamic and electrophysiological responses to different stimulus contrasts. (A) Differential orientation maps representing responses to vertical (dark regions) and horizontal gratings (bright regions) at low-contrast (left) and high-contrast (right) stimulation recorded at 610 nm. For these maps, data were averaged over five recording blocks. In both maps, gray levels correspond to the same activity levels. The rightmost image depicts the recorded area. Black and white arrowheads in the activity maps indicate corresponding regions. In the recorded area, lateral (lat) is toward the top and anterior (ant) is toward the right. Scale bar, 1 mm. (B) Time course of the hemodynamic responses at 610 nm averaged over 26 recording blocks for each contrast level. (C) Peristimulus time histograms calculated from the normalized spike rate averaged over all recording blocks and sites (26 blocks, four recording sites). Error bars in (B) and (C) show means \pm SD. (D) Power spectra of all trials ($n = 26$ recording blocks) sorted in an ascending order according to the hemodynamic response strength at 610 nm for each contrast level. The power of each frequency bin was renormalized to the range between the maximum and minimum power observed throughout all trials and over both contrast levels in the respective frequency bin. A low-pass filter was applied on every single frequency bin of the resulting plot (kernel size includes 10 trials). (E and F) Correlation analysis between hemodynamic responses and spike rates (E) and hemodynamic responses and gamma power in the upper frequency range (F), calculated on the basis of individual data blocks of eight stimulus presentations summarizing data from both contrast levels. Each point reflects the averaged response over eight repetitions of the presentation of the eight different stimuli. Therefore, each point reflects the average of 64 responses measured over a period of about 19 min.

LFPs depends not only on the number of active synapses within a cortical volume but to a crucial extent on the synchronicity of the

synaptic events (8). Because of the short duration of synaptic currents, they summate most effectively if synchronized with a preci-

sion in the millisecond range. The precision of synchronous firing is in turn positively correlated with the frequency at which synchro-

Fig. 2. Spontaneous variations of hemodynamic and electrophysiological response under conditions of constant stimulus intensity. (A) Time course of hemodynamic signals at 610 nm averaged over the respective 33.3% of trials with strong (red line), medium (blue line), and weak (green line) hemodynamic responses ($n = 26$ recording blocks for each curve). (B) PSTHs of the spike rates for the same trials as depicted by (A). Error bars in (A) and (B) show means \pm SD. (C) Power spectra of all trials ($n = 78$ recording blocks) sorted in an ascending order according to hemodynamic response strength at 610 nm. The observed minimum and maximum activity is coded by black and yellow, respectively. A low-pass filter was applied on each single frequency bin of the resulting plot (kernel size includes 10 trials).

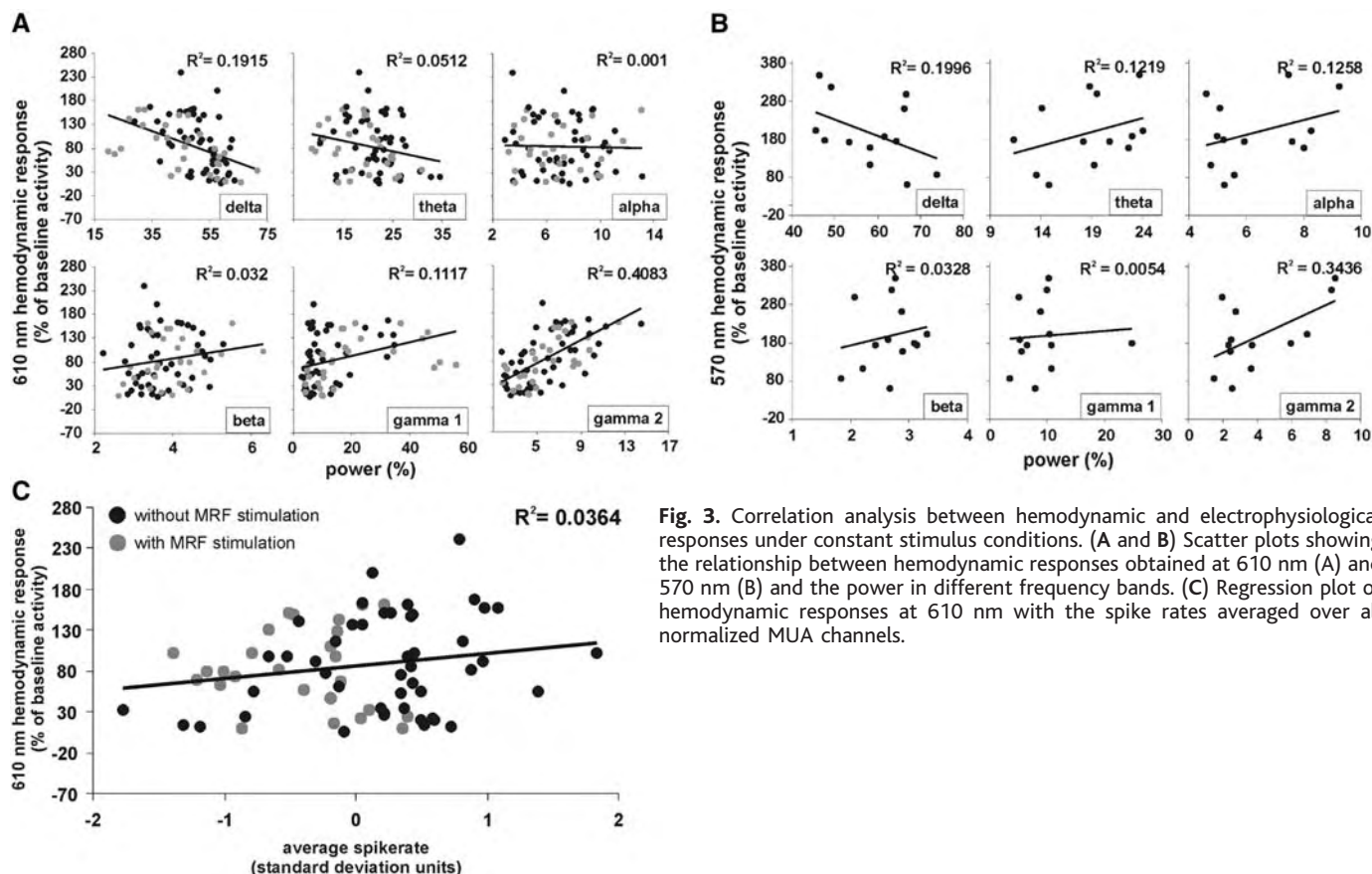
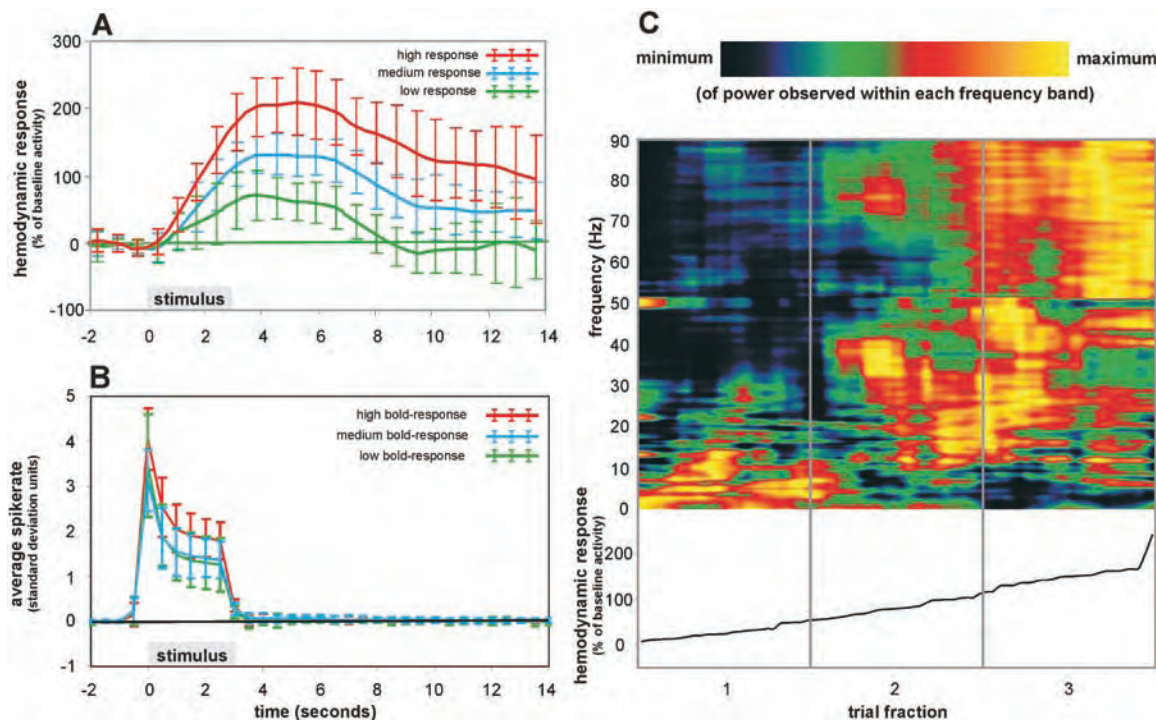


Fig. 3. Correlation analysis between hemodynamic and electrophysiological responses under constant stimulus conditions. (A and B) Scatter plots showing the relationship between hemodynamic responses obtained at 610 nm (A) and 570 nm (B) and the power in different frequency bands. (C) Regression plot of hemodynamic responses at 610 nm with the spike rates averaged over all normalized MUA channels.

nized groups of neurons oscillate (15, 22) and reaches a maximum when cells engage in high-frequency gamma oscillations.

How do changes in the temporal patterning of responses influence hemodynamic responses? There is evidence that synchronization in the gamma frequency range is associated with oscillatory, tightly synchronized discharges of inhibitory interneurons (23) leading to periodic inhibition of pyramidal cells. These inhibitory postsynaptic potentials, in turn, synchronize pyramidal cells by confining their discharges to the depolarizing peaks of the membrane potential oscillations (24). Thus, when cortical networks engage in gamma oscillations, inhibitory interneurons are highly active, and as their discharges are phase-locked to the oscillations (23), their activity increases with oscillation frequency. Therefore, we propose that the hemodynamic responses associated with gamma oscillations are mainly initiated by the firing of inhibitory interneurons.

Although inhibitory interneurons constitute only about 20% of the cortical neurons, it is likely that they substantially contribute to local energy consumption. They fire at very high frequencies and distribute their numerous synapses exclusively within adjacent cortical volumes. The hypothesis that interneuron activity is a major cause of the hemodynamic response (25) is well compatible with the fact that interneurons contain enzymes for the synthesis of vasoactive compounds such as NO and vasoactive peptides (26, 27). In this cascade, the elevated calcium concentration is suggested to play a major role (25).

This interpretation resolves some of the discrepancies between BOLD studies and unit recordings. Deficits in visual processing in amblyopia are well reflected by evoked potentials and hemodynamic responses, although they are undetectable in discharge rates (7). Attentional shifts in the absence of sensory stimulation (28) and mental imagery (29) are associated with BOLD responses, and these cognitive processes are associated with increased oscillatory activity in the gamma frequency band (14, 30). Other cognitive and executive functions such as figure-ground segmentation, expectancy, sensory-motor coordination, short-term memory, and movement preparation are associated with enhanced oscillatory activity in the beta and gamma frequency range (31, 32). Hemodynamic responses may thus be ideally suited to visualize neural processes associated with higher cognitive and executive functions.

References and Notes

1. R. D. Frostig, E. E. Lieke, D. Y. Ts'o, A. Grinvald, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6082 (1990).
2. S. Ogawa, T. M. Lee, A. R. Kay, D. W. Tank, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9868 (1990).
3. G. Rees, K. Friston, C. Koch, *Nat. Neurosci.* **3**, 716 (2000).
4. A. J. Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10765 (2002).

5. O. J. Arthurs, E. J. Williams, T. A. Carpenter, J. D. Pickard, S. J. Boniface, *Neuroscience* **101**, 803 (2000).
6. S. Sheth et al., *Neuroimage* **19**, 884 (2003).
7. K. E. Schmidt, W. Singer, R. A. Galuske, *J. Neurophysiol.* **91**, 1661 (2004).
8. U. Mitzdorf, W. Singer, *Exp. Brain Res.* **33**, 371 (1978).
9. N. K. Logothetis, J. Pauls, M. Augath, T. Trinath, A. Oeltermann, *Nature* **412**, 150 (2001).
10. C. Kayser, M. Kim, K. Ugurbil, D. S. Kim, P. Konig, *Cereb. Cortex* **14**, 881 (2004).
11. C. Mathiesen, K. Caesar, N. Akgoren, M. Lauritzen, *J. Physiol.* **512**, 555 (1998).
12. D. A. McCormick, T. Bal, *Annu. Rev. Neurosci.* **20**, 185 (1997).
13. M. Steriade, *Science* **272**, 225 (1996).
14. P. Fries, J. H. Reynolds, A. E. Rorie, R. Desimone, *Science* **291**, 1560 (2001).
15. S. Herculano-Houzel, M. H. Munk, S. Neuenschwander, W. Singer, *J. Neurosci.* **19**, 3992 (1999).
16. R. Rodriguez, U. Kallenbach, W. Singer, M. H. Munk, *J. Neurosci.* **24**, 10369 (2004).
17. Materials and methods are available as supporting material on Science Online.
18. G. M. Boynton, S. A. Engel, G. H. Glover, D. J. Heeger, *J. Neurosci.* **16**, 4207 (1996).
19. O. J. Arthurs, S. Boniface, *Trends Neurosci.* **25**, 27 (2002).
20. N. Pouratian et al., *Magn. Reson. Med.* **47**, 766 (2002).
21. D. Attwell, S. B. Laughlin, *J. Cereb. Blood Flow Metab.* **21**, 1133 (2001).
22. M. Volgushev, M. Chistiakova, W. Singer, *Neuroscience* **83**, 15 (1998).
23. R. D. Traub, M. A. Whittington, I. M. Stanford, J. G. Jefferys, *Nature* **383**, 621 (1996).
24. R. Azouz, C. M. Gray, *J. Neurosci.* **19**, 2209 (1999).
25. M. Lauritzen, *Nat. Rev. Neurosci.* **6**, 77 (2005).
26. J. G. Valtchanoff et al., *Neurosci. Lett.* **157**, 157 (1993).
27. B. Cauli et al., *J. Neurosci.* **24**, 8940 (2004).
28. S. Kastner, M. A. Pinsk, P. De Weerd, R. Desimone, L. G. Ungerleider, *Neuron* **22**, 751 (1999).
29. S. M. Kosslyn, G. Ganis, W. L. Thompson, *Nat. Rev. Neurosci.* **2**, 635 (2001).
30. M. M. Muller, T. Gruber, A. Keil, *Int. J. Psychophysiol.* **38**, 283 (2000).
31. C. Tallon-Baudry, *J. Physiol. (Paris)* **97**, 355 (2003).
32. A. K. Engel, P. Fries, W. Singer, *Nat. Rev. Neurosci.* **2**, 704 (2001).
33. We thank H. Klon-Lipok and E. Scheibinger for technical assistance and M. Wibrall, M. Lutzenburg, and G. Pipa for helpful comments and suggestions on the project.

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5736/948/DC1

Materials and Methods

Figs. S1 to S4

References

10 February 2005; accepted 10 June 2005

10.1126/science.1110948

Coupling Between Neuronal Firing, Field Potentials, and fMRI in Human Auditory Cortex

Roy Mukamel,¹ Hagar Gelbard,¹ Amos Arieli,¹ Uri Hasson,² Itzhak Fried,^{3,4*} Rafael Malach^{1*}

Functional magnetic resonance imaging (fMRI) is an important tool for investigating human brain function, but the relationship between the hemodynamically based fMRI signals in the human brain and the underlying neuronal activity is unclear. We recorded single unit activity and local field potentials in auditory cortex of two neurosurgical patients and compared them with the fMRI signals of 11 healthy subjects during presentation of an identical movie segment. The predicted fMRI signals derived from single units and the measured fMRI signals from auditory cortex showed a highly significant correlation ($r = 0.75$, $P < 10^{-47}$). Thus, fMRI signals can provide a reliable measure of the firing rate of human cortical neurons.

A major concern in the rapidly expanding field of functional magnetic resonance imaging (fMRI) has been the absence of a quantitative relationship between blood oxygenated level-dependent (BOLD) fMRI signals and neuronal activity. Several studies have attempted to characterize this relationship (1–10). In anesthetized monkeys, a higher correlation of the

fMRI signal to the local field potential (LFP) was found compared with spike activity (11). However, the implications of these studies to the awake, conscious human brain are unclear.

Recently, we reported that movie stimuli are particularly effective in producing a widespread and robust correlation in the evoked fMRI signals across different subjects (12). Here, we used this phenomenon of intersubject correlation to examine the nature of the coupling between fMRI signals and neuronal activity in the sensory cortex of alert humans.

We recorded from 53 single neurons in Heschl's gyrus (auditory cortex) of two native English-speaking patients with epilepsy monitored with intracranial depth electrodes for potential surgical treatment (13). Recordings were done while the patients saw two repetitions of a 9-min segment from a popular

¹Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel. ²Center for Neural Science, New York University, New York, NY 10003, USA.

³Division of Neurosurgery, David Geffen School of Medicine, and Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles (UCLA), Los Angeles, CA 90095, USA. ⁴Functional Neurosurgery Unit, Tel Aviv Medical Center and Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel.

*To whom correspondence should be addressed. E-mail: rafi.malach@weizmann.ac.il (R.M.); ifried@mednet.ucla.edu (I.F.)

English-speaking movie. Twenty out of 30 neurons in patient 1 and 17 out of 23 in patient 2 showed reproducible auditory responses. The spiking activity of these neurons in each patient was summed and converted into a predicted fMRI BOLD response (spike predictor) by convolution with a standard hemo-

dynamic response function (gamma function) (2, 14) (Fig. 1). In parallel, we performed fMRI scans on 11 healthy subjects while they watched the same movie segment (13). All subjects spoke English as a second language. We then used the spike predictor of each patient (e.g., Fig. 1C for patient 1) as a

regressor in a conventional general linear model (GLM) analysis of the fMRI signals obtained from the healthy subjects. Figure 2 shows the multiparticipant average activation map of 11 subjects as derived from the spike predictor of patient 1, projected onto the reconstructed folded hemispheres of the patient. The map has its basis in the average of all single-subject data, transformed into common Talairach space (15). On the basis of postoperative computer tomography scans, we could estimate the location of the recording electrode ("e" in Fig. 2). The most significant activation foci were localized to Heschl's gyrus in close proximity to the electrode's location. Significant activation in the same region was also found in individual subjects (see individual maps in fig. S1). In the second patient, the data followed a similar trend (fig. S2). No significant activation was revealed with the use of a predictor derived from the soundwave amplitude (13). Furthermore, control fMRI data, recorded while subjects viewed the same movie segment with no sound or viewed a different movie segment, revealed no correlation with the spike predictor (fig. S3).

We sampled the fMRI time course in Heschl's gyrus of each subject by using the spike predictors of the patients. To ensure statistical independence of region of interest (ROI) definition and sampling of signal time course, we used the internal localizer design in which nonoverlapping time segments of the spike predictor were used for ROI definition and for sampling (13, 16). Time courses were then averaged across subjects. The average fMRI time courses are shown in the bottom graphs of Fig. 3 and in fig. S4. These time courses are superimposed on the spike predictor (black trace). The correlation coefficients between the averaged fMRI signal and the spike predictors were 0.75 ($P < 10^{-47}$) for patient 1 and 0.56 ($P < 10^{-28}$) for patient 2. The sampled fMRI signals were slightly different for the two patients, because their

Fig. 1. The fMRI BOLD predictor created from spikes of patient 1. (A) Raster display of spike trains from all 20 cells during the first (blue) and second (red) movie presentation. The black histogram displays the population summed activity. The summed activity was convolved with a gamma hemodynamic response function (B). The resulting time course was Z-score normalized (C) and used as a predictor for the BOLD fMRI signal. a.u., arbitrary units.

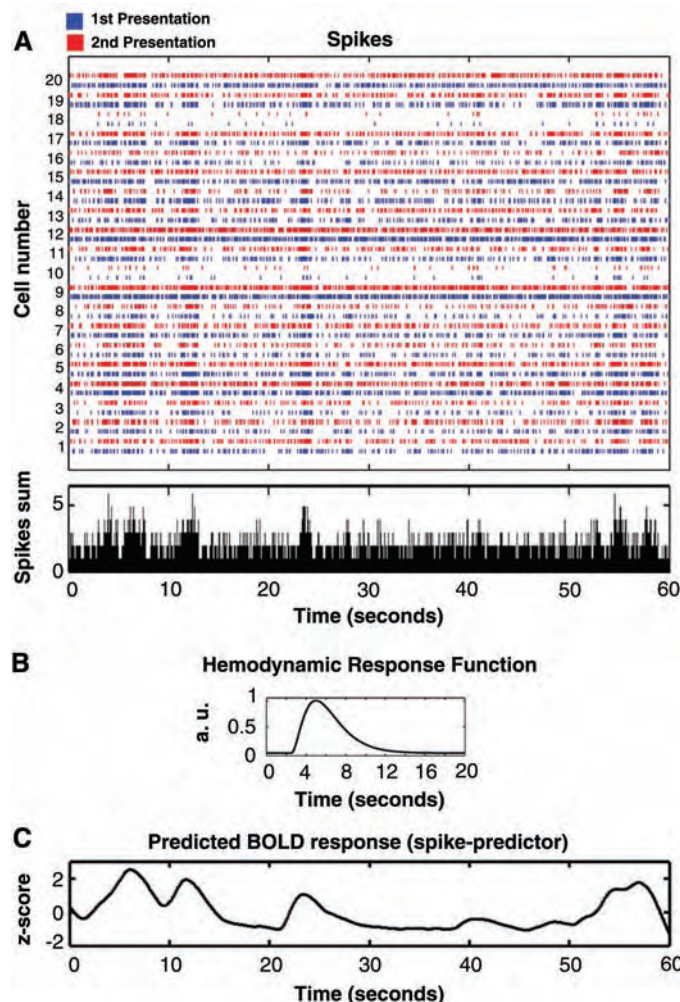
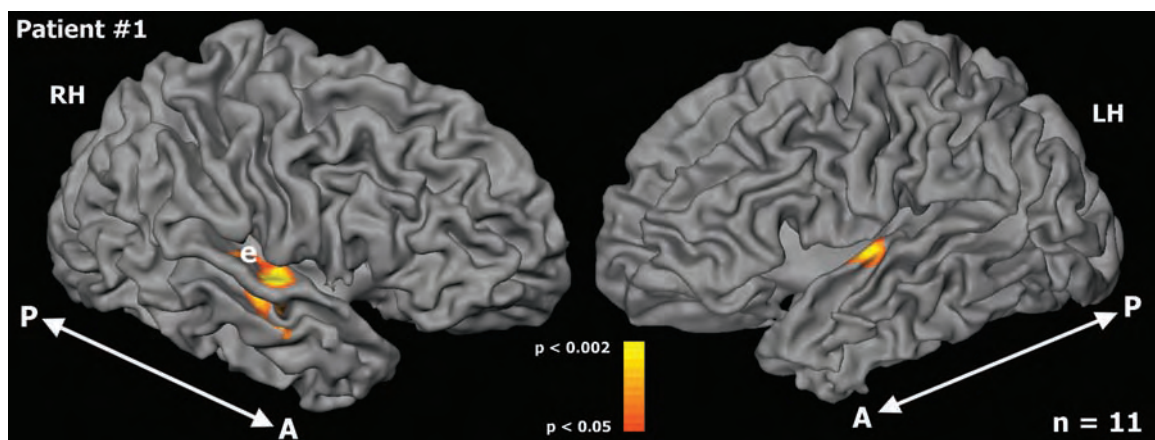


Fig. 2. The fMRI BOLD activation map derived from the spike predictor. Multisubject map ($n = 11$) obtained with use of the spike predictor of patient 1. The fMRI activations are shown on the reconstructed folded hemispheres of the patient. Activations were obtained with a random effect multi-GLM test by using the predictor obtained from the average spike train (20 cells) of patient 1. Note the localized activations in Heschl's gyrus corresponding to the site of spike recordings in the patient (denoted by "e"). RH and LH, right and left hemisphere, respectively; A, anterior; P, posterior.



spike predictors defined slightly different ROIs, possibly due to the fact that the electrode position in the two patients was not identical [for exact Talairach coordinates see (13)].

The amount of correlation between the fMRI signals and spike predictors modulated along the 9-min movie timeline. For example, during stretches of silence in the movie there was no stimulus-evoked activity in auditory cortex, and no correlation is expected between the spontaneous neural activity in different individuals. In order to isolate times of high correlation between different individuals, we calculated a moving correlation function between the fMRI signals of the healthy subjects (green traces in Fig. 3) (13). Periods of high intersubject correlation (light gray shaded regions) largely coincided with periods of high correlation between the BOLD fMRI signal and the spike predictors. Thus, although the correlations between the neuronal and fMRI signals were 0.75 and 0.56 for the two patients for the entire duration of the recording, these correlations climbed to 0.9 ($P < 10^{-30}$) and 0.72 ($P < 10^{-19}$) when only 30% of the time

course, showing the highest interparticipant correlations, was considered.

The electrophysiological recordings from the two patients were continuous, spanning 1 to 14 kHz, allowing us to examine also the relationship between the LFPs and the fMRI signals. LFPs are typically of lower frequencies (~ 1 to 130 Hz) and likely relate to dendritic synaptic activity (1). We created fMRI predicted signals from the power modulations of the LFP signals at different frequencies (LFP predictors) (13). Figure 4A depicts the correlation between the different LFP predictors (corresponding to different frequency bands) and the spike predictor (cyan) and also their correlation with the average fMRI signal sampled with the spike predictor (orange) for patient 1. Interestingly, at different frequency bands there was a clear inversion of the correlation value between the fMRI data and the LFP signal from negative correlation at low LFP frequencies (5 to 15 Hz, green) to strong positive correlation at high frequencies (40 to 130 Hz, yellow). Also there was a strong correlation between the spike predictor and LFP predictor at high frequencies (cyan

trace). We generated LFP predictors for the two frequency bands (5 to 15 Hz and 40 to 130 Hz) (13). All three predictors (spikes and low- and high-frequency LFPs) were most significantly correlated with the fMRI signal in Heschl's gyrus (with a sign inversion for the low frequency LFP predictor) (Fig. 4, B to D).

Correlation values of high-frequency LFP and spike predictors with fMRI BOLD signal of each subject were compared. The results show similar correlations derived from these two predictors: for patient 1, spike predictor correlation of 0.5 ± 0.04 SEM and LFP predictor of 0.46 ± 0.04 ; for patient 2, spike predictor correlation of 0.39 ± 0.04 SEM and LFP predictor correlation of 0.39 ± 0.04 . Thus, in patient 1, the correlation with BOLD was slightly better for the spike predictor than the high-frequency LFP-predictor ($P < 0.05$, paired t test), whereas no significant difference was found in patient 2 ($P < 0.48$, paired t test). For spike and LFP predictors, their corresponding fMRI signals, and spectral analysis of stimulus soundwave, see fig. S5.

Lastly, although the measured time courses were rather short, the availability of both spiking activity and fMRI responses to the same stimulus allowed us to obtain a data-driven, objective estimate of the human hemodynamic response function, coupling the spikes and fMRI BOLD signal. By using standard deconvolution methods (13), we obtained such functions for the two fMRI data sets, sampled at a rate of 1 and 1/3 Hz (orange and purple traces in fig. S6). The obtained functions were quite compatible with the current estimates of the BOLD hemodynamic response function (black trace) (2, 14). Noticeably, the data sampled at the faster rate exhibited a small but significant ($P < 10^{-4}$) initial decrease in the BOLD signal of four out of six subjects (fig. S6, gray circles) (13, 17, 18).

Our results indicate a high linear correlation between spiking activity, high-frequency LFP, and fMRI BOLD signal measured in human auditory cortex during natural stimulation. We cannot rule out the possibility that under different physiological conditions or in different brain regions there may be a decoupling between spiking activity and the BOLD signal (11). Because in our stimulation paradigm the spiking activity was highly correlated with the high-frequency LFP (cyan trace, Fig. 4A), our results cannot identify one or the other as the driving source behind the BOLD signal (9–11, 19, 20). However, regardless of the mechanism underlying the BOLD signal, the broad methodological implication of our findings lies in the demonstration that, at least under natural stimulus conditions, BOLD fMRI signals can be trusted as a faithful measure of the average firing rate of the underlying population.

The high correlation found does not necessarily imply lack of variability in the fMRI

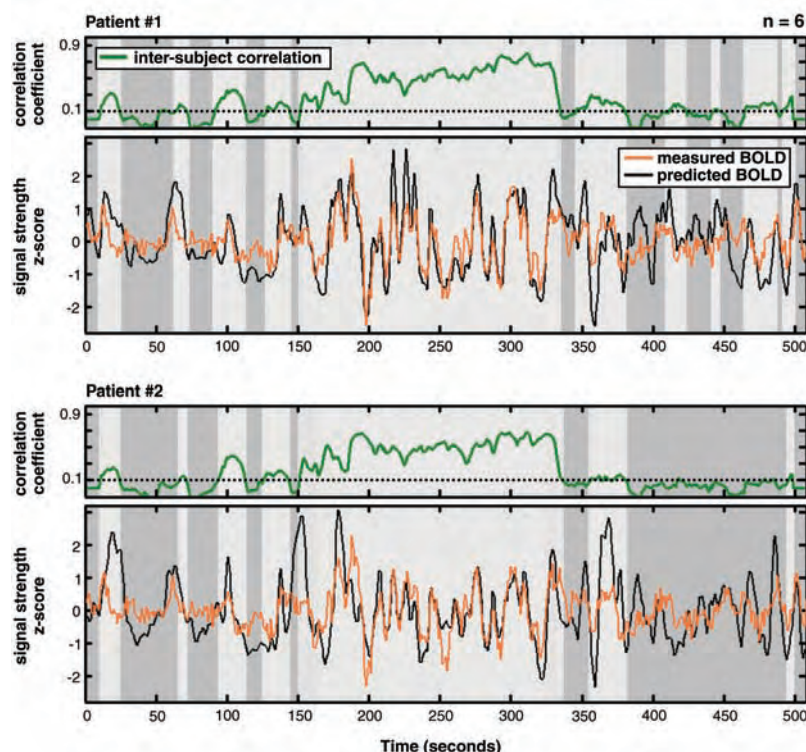


Fig. 3. Correlation of the spike predictor with measured fMRI activation. fMRI BOLD activations in Heschl's gyrus of each of the six subjects were sampled by using the spike predictor of each patient (top, patient 1; bottom, patient 2). Bottom graphs depict the average measured BOLD activation of all six subjects (orange traces) together with the predicted BOLD created from the average spike train of each patient (black traces). The overall correlations between the average BOLD signal and the spike predictors were 0.73 for patient 1 and 0.55 for patient 2. Top graphs depict the average intersubject correlation (green traces) between BOLD activations of all subject pairs calculated in sliding windows of 21 s (13). Light gray regions correspond to times in which the average intersubject correlation was greater than 0.1, whereas dark gray regions correspond to times in which the average intersubject correlation was smaller than 0.1. The correlation between the BOLD signal and spike predictors increased at high (> 0.1) levels of intersubject correlations to 0.81 (0.6) (light gray regions) and further increased to 0.9 (0.72) in the top 30% epochs.

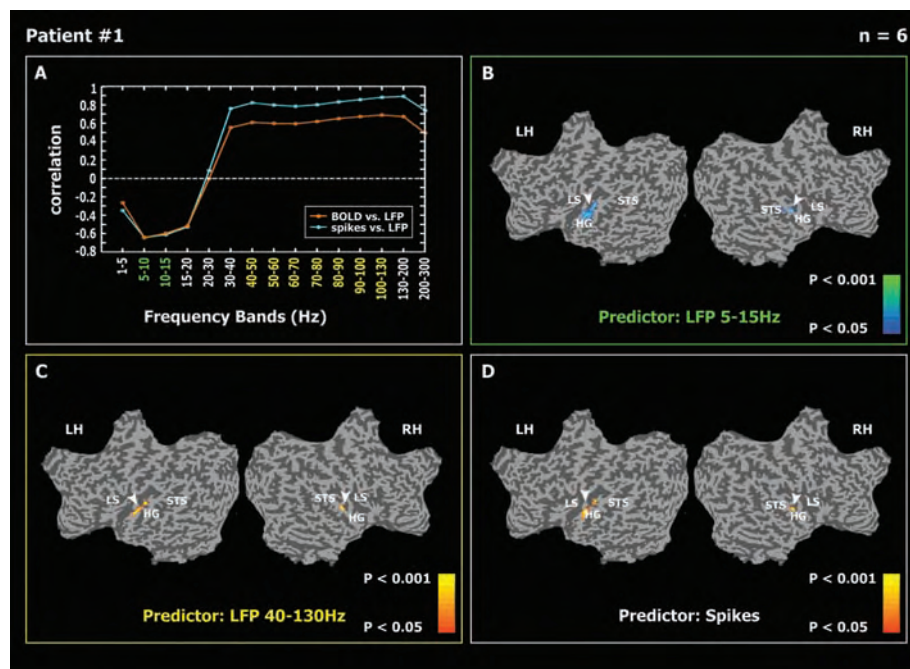


Fig. 4. Correlation maps between fMRI BOLD signal and different components of the electro-physiological measurements. (A) The correlation of the different LFP predictors of patient 1 (13) with the average fMRI BOLD signal in Heschl's gyrus (orange trace) and with the spike predictor (cyan trace) as a function of frequency band. Note the strong negative correlations between the BOLD activation and the low-frequency LFPs (5 to 15 Hz) and the strong positive correlation with the high-frequency LFPs (40 to 130 Hz). (B to D) The multisubject random-effect GLM map correlating the BOLD signal of six participants with the low-frequency (5 to 15 Hz) LFP predictor (B), the high-frequency (40 to 130 Hz) LFP predictor (C), and the spike predictor (D). LS, lateral sulcus; STS, superior temporal sulcus; HG, Heschl's gyrus; RH and LH, right and left hemisphere, respectively. Arrowheads point to regions of highly significant correlation in Heschl's gyrus.

signals. Note that the spikes-to-BOLD correlation obtained for group-averaged fMRI responses was higher (0.75 for patient 1 and 0.56 for patient 2) than that obtained for individual subjects (0.5 and 0.46, respectively).

A particularly important result was the inversion in the coupling of LFP to BOLD, moving from positive correlations at high, gamma-range frequencies (40 to 130 Hz) (21) to negative correlations at the low, alpha-range frequencies (5 to 15 Hz) (22, 23). This result is compatible with the notion that electroencephalogram signals in the alpha band frequencies are diminished in times of high cognitive processing and that sensory stimulation leads to desynchronization of slow potentials (24).

The fact that the correlation between the spikes, high-frequency LFPs, and fMRI signal in Heschl's gyrus was high despite the extremely noisy MRI environment indicates that human auditory representations at the cortex level deemphasize low-level auditory signals such as loud beeps (25, 26). This is also illustrated in the relatively low (albeit still significant) fMRI responses to high sound-amplitude events (e.g., gunshots and explosions) during the movie compared with speech

events (fig. S4) and the weak correlation of the activations with the sound-wave amplitude (figs. S3 and S5).

Our results also have implications for human sensory representations during natural stimulation. The single-unit signal was derived from a small population (~20) of active neurons, whereas the fMRI signals reflect the population activity of millions of neurons (8). The fact that firing rate in a mere ~20 responsive neurons was sufficient to capture such a large fraction of the fMRI signal variance points to a highly distributed profile (27) of slow auditory responses. In such a representation, auditory neurons contain a strong element of correlation between neighboring neurons during natural sensory stimulation, because they are all driven, albeit to a varying extent, by multiple auditory events (Fig. 1A).

Several intervening factors could have disrupted the correlation between the electro-physiological and fMRI signals. First, the electrophysiological signals were recorded in a quiet hospital room, whereas the fMRI signals were recorded in a typical, acoustically noisy fMRI setup. Second, the small sample of neurons we recorded was unlikely to precisely

represent the averaged response profile of the neuronal populations in the imaged fMRI voxels. Lastly, there were individual differences between the patients and the fMRI subjects. Thus, under optimal conditions, such as recording both signals in the same participant, the real physiological coupling between the neuronal firing rates and fMRI BOLD signals may be even stronger than measured.

References and Notes

- N. K. Logothetis, B. A. Wandell, *Annu. Rev. Physiol.* **66**, 735 (2004).
- G. M. Boynton, S. A. Engel, G. H. Glover, D. J. Heeger, *J. Neurosci.* **16**, 4207 (1996).
- A. M. Dale, R. L. Buckner, *Hum. Brain Mapp.* **5**, 329 (1997).
- G. K. Aguirre, E. Zarahn, M. D'Esposito, *Neuroimage* **8**, 360 (1998).
- D. A. Handwerker, J. M. Ollinger, M. D'Esposito, *Neuroimage* **21**, 1639 (2004).
- D. J. Heeger, A. C. Huk, W. S. Geisler, D. G. Albrecht, *Nat. Neurosci.* **3**, 631 (2000).
- G. Rees, K. Friston, C. Koch, *Nat. Neurosci.* **3**, 716 (2000).
- I. Levy, U. Hasson, R. Malach, *Curr. Biol.* **14**, 996 (2004).
- D. Attwell, C. Iadecola, *Trends Neurosci.* **25**, 621 (2002).
- C. Iadecola, *Nat. Rev. Neurosci.* **5**, 347 (2004).
- N. K. Logothetis, J. Pauls, M. Augath, T. Trinath, A. Oeltermann, *Nature* **412**, 150 (2001).
- U. Hasson, Y. Nir, I. Levy, G. Fuhrmann, R. Malach, *Science* **303**, 1634 (2004).
- Materials and methods are available as supporting material on Science Online.
- R. Goebel, *Neuroimage* **3**, S604 (1996).
- J. Talairach, P. Tournoux, *Co-Planar Stereotaxic Atlas of the Human Brain* (Thieme Medical, New York, 1988).
- Y. Lerner, T. Hendler, R. Malach, *Cereb. Cortex* **12**, 163 (2002).
- B. M. Ances, *J. Cereb. Blood Flow Metab.* **24**, 1 (2004).
- E. Yacoub, X. Hu, *Magn. Reson. Med.* **45**, 184 (2001).
- K. Caesar, K. Thomsen, M. Lauritzen, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 16000 (2003).
- C. Mathiesen, K. Caesar, N. Akgoren, M. Lauritzen, *J. Physiol.* **512**, 555 (1998).
- M. Siegel, P. König, *J. Neurosci.* **23**, 4251 (2003).
- M. Moosmann et al., *Neuroimage* **20**, 145 (2003).
- H. Laufs et al., *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11053 (2003).
- T. V. Searwards, M. A. Searwards, *Int. J. Psychophysiol.* **32**, 35 (1999).
- P. Belin, R. J. Zatorre, P. Lafaille, P. Ahad, B. Pike, *Nature* **403**, 309 (2000).
- A. L. Giraud et al., *J. Neurophysiol.* **84**, 1588 (2000).
- J. V. Haxby et al., *Science* **293**, 2425 (2001).
- This study was funded by an Israeli Science Foundation grant to R.M., a U.S.-Israel Binational Science Foundation grant to I.F. and R.M., and a National Institute of Neurological Disorders and Stroke grant to I.F. The authors thank the patients for their cooperation in participating in the study. We also thank Y. Nir for helpful suggestions; M. Harel for help with the fMRI data; D. Malah for help with the deconvolution algorithm; D. Heeger, S. Gilaie-Dotan, and A. Amedi for reading the manuscript; and E. Isham, E. Ho, T. A. Fields, E. Behnke and C. Wilson for technical assistance.

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5736/951/DC1

Materials and Methods

Figs. S1 to S6

10 February 2005; accepted 10 June 2005

10.1126/science.1110913

Automated Protein Purification

The BioSprint 15 Ni-NTA and Strep-Tactin kits bring automated protein purification within reach of more researchers. BioSprint 15 protein kits save time and effort by providing fast purification of high-purity histidine- and streptavidin-tagged recombinant proteins, ready to use in any downstream application. The easy-to-use BioSprint 15 platform is a stand-alone instrument with a small footprint that enables automated parallel purification of up to 15 high-purity protein samples from prokaryotic or eukaryotic cell lysates. By using magnetic-bead technology, BioSprint 15 kits provide a rapid procedure that can purify up to 100 µg protein per sample in less than 45 min. Since the workstation transfers magnetic beads instead of liquids, it uses minimal amounts of reagents.

Qiagen For information 800-426-8157 www.qiagen.com

Cell Proliferation Analysis

The Guava CellGrowth assay, when used in conjunction with the Guava EasyCyte System, provides cell proliferation analysis at the single-cell level, offering cell proliferation assessments without the use of more complex fluorescence methods out to five generations. In addition, the assay offers a broader linear range and is more quantitative than colorimetric methods. The CellGrowth Assay makes use of a well-characterized cell tracking dye that diffuses freely into cells and is retained within the cell without affecting cellular function. The dye is not transferred to adjacent cells. The assay makes use of two dyes, a cell-permanent painting dye and a cell-impermanent DNA-binding dye, to distinguish live or dead proliferated cells from live or dead resting cells. For each round of cell division, the relative fluorescence intensity of the dye is decreased by half. The assay is suitable for monitoring the growth of primary cultures, such as peripheral blood mononuclear cells, activated T or B cells, peripheral blood lymphocytes, thymocytes, hematopoietic cells, splenocytes, and fibroblasts.

Guava Technologies For information 44-208-546-0869
www.guavatechnologies.com

Mutation Detection Kit

The Surveyor Mutation Detection Kit for standard gel electrophoresis contains the key component Surveyor Nuclease, which recognizes and cleaves DNA fragments at any type of mismatch site. The Surveyor Nuclease can be used to detect all single-base substitutions and small insertions and deletions; find mutations in polymerase chain reaction (PCR) products 0.2 to 3.5 kb in size; reveal the number of mutations in a specific PCR product; show whether 1, 2, or 3 mutations are present; and discover mutations in pooled genomic DNA samples or pooled PCR products representing up to 16 alleles.

Transgenomic For information 402-452-5452 www.transgenomic.com

Advanced Biomarker Discovery

The ClinProt system for advanced biomarker discovery, identification, and validation in clinical proteomics has expanded capabilities. The extensive ClinProt portfolio of magnetic beads for sample pre-fractionation has been expanded with novel capabilities to address

biomarker analysis of large proteins, targeted glycoprotein analysis, and antigen-specific enrichment of selected biomarker molecules. In addition to these beads, a new ClinProTools 2.0 bioinformatics package offers streamlined data interpretation. It combines intuitive data visualization with sophisticated pattern recognition algorithms for the determination of predictive biomarker panels.

Bruker Daltonics For information 978-667-9580 www.bruker-biosciences.com

Heating Circulators

The 8200 Series is a line of large capacity heating circulators designed to provide precise temperature control for a diverse range of applications. The series offers a choice of three different micro-processor-based controllers, integrated simplex or duplex pump, and a 28-l reservoir. The units feature an operating temperature range of ambient +5°C to 150°C, excellent temperature stability, and a large 12 1/8-inch by 10 3/8-inch reservoir opening. The top-of-the-line 8212 Circulator features a programmable controller with bright LCD display, ±0.01°C temperature stability, multi-language help menus, and sophisticated time and temperature programming. It can store up to ten 50-step programs and offers extensive data-logging capability, including communications support for Microsoft Excel, National Instruments LabView, and the Palm OS. Other standard features include a five-speed pump, remote probe control capability, and a built-in RS-232 interface.



PolyScience For information 800-229-7569 www.polyscience.com

Literature

Fluorescent Proteins 2005 is a 20-page guide describing Evrogen's novel fluorescent proteins for biological research and commercial

biotechnology applications. The guide includes basic fluorescent proteins such as TurboGFP (a bright and fast-maturing green fluorescent protein), PhyYFP and PhiYFP-m (true yellow fluorescent proteins), t-HcRed (a non-oligomeric far-red fluorescent protein), and the red fluorescent protein JRed. It also features corresponding expression vectors, recombinant proteins, and protein antibodies as well as proteins that change color and intensity in response to light of certain wavelengths. This reference discusses the general properties of these proteins and details a variety of in vivo applications, including gene expression analysis,

protein localization, cell labeling, and protein-protein interaction and co-localization studies.

Axxora For information 800-550-8825 www.axxora.com

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and government organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier by visiting www.science.labvelocity.com on the Web, where you can request that the information be sent to you by e-mail, fax, mail, or telephone.

For more information visit **GetInfo**,
Science's new online product index at
<http://science.labvelocity.com>

From the pages of GetInfo, you can:

- Quickly find and request free information on products and services found in the pages of *Science*.
- Ask vendors to contact you with more information.
- Link directly to vendors' Web sites.